

PHYSICAL CHEMISTRY

PHYSICAL CHEMISTRY

BY

J. N. BRÖNSTED

Professor in Copenhagen University

Translated from the Danish Edition 1936 by
R. P. Bell, M.A., B.Sc., Fellow of Balliol
College, Oxford, with a Foreword by F. G.
Donnan, D.Sc., F.R.S., Professor of Chemistry
in the University of London.



WILLIAM HEINEMANN LTD.

99 GREAT RUSSELL STREET

LONDON, W.C.1.

SET IN MODERN EXTENDED TYPE
AND PRINTED IN GREAT BRITAIN
BY THE WHITEFRIARS PRESS LTD.,
LONDON AND TONBRIDGE

First Published November, 1937

FOREWORD

By F. G. DONNAN, C.B.E., LL.D., D.Sc., F.R.S.,
Professor of Chemistry in the University of London.

It is both a pleasure and a privilege to have this opportunity of welcoming an English translation of Professor Brönsted's excellent and most valuable Text-book of Physical Chemistry. An outstanding feature of this work lies in its systematic application of the thermodynamical methods of Willard Gibbs. The clarity, security and elegance of treatment thereby obtained are apparent on almost every page. For this achievement, Professor Brönsted is to be heartily congratulated.

As we should expect from the author, the chapters on solutions of electrolytes are of fundamental importance, and will be of the greatest value to all students of chemistry and physics. The sections dealing with the generalised theory of acids and bases and the related phenomena of protolysis and protolytic equilibria will serve to familiarise English students with very important advances due to the Danish school of physical chemistry.

In the chapter on Electrochemistry, the employment of Gibbs' chemical potential in the modified form of Guggenheim's "electrochemical potential" and the author's "electromotive potential" makes the treatment of a large part of this subject both elegant and exact.

These are only a few samples of the many good things contained in this book. Although the main treatment is thermodynamical, excellent discussions from the point of view of kinetics and molecular theory are interspersed throughout the text, and there is also a valuable chapter on reaction kinetics (including photochemical reactions).

The work of translation has been well and faithfully done by Mr. R. P. Bell. The text is as smooth and clear as one written originally in English. It is obvious that the translator has a complete command of the Danish language, whilst his name is a sufficient guarantee for the scientific accuracy of the English text.

It is not only a privilege but a duty to recommend this book to all students of physical chemistry. Coming from the famous laboratory directed by Professor Brönsted in Copenhagen, it is destined to occupy a unique and pre-eminent position in the advancement of physico-chemical science.

PREFACE TO THE ENGLISH EDITION

THE present book is a translation of the author's "*Lærebog i fysisk Kemi*," which was published in 1936, and forms the basis for instruction in physical chemistry at Copenhagen University and the Danish Technical College.

In this text-book special emphasis is laid on the thermodynamic foundation of physico-chemical laws. The treatment employed is based upon the "thermodynamic functions" of Gibbs, since in the author's opinion this is the easiest way of attaining the combination of simplicity and rigour which should be the aim of a theoretical text-book. It is often supposed that the conception of thermodynamic functions is a difficult one, and that their use demands special powers of abstract thought. It is the author's view, however, that this is not in general the case, while even if such difficulties should occur they will be amply compensated by the uniformity of treatment and certainty of operation which characterise the methods of Gibbs.

In dealing with the basis of the laws of thermodynamics, the presentation in this book differs considerably from that commonly employed. The types of work corresponding to the different forms of energy are introduced at the beginning of the treatment, entropy appearing in a simple manner as the extensive factor of thermal work. This method of treatment leads to a modified conception of the relation between heat and work, and hence to an interpretation of the working of heat engines which is essentially different to that given by Clausius.

In addition to the thermodynamic aspect the subject-matter is also dealt with from the point of view of kinetic and molecular theory, partly in separate sections and partly interpolated in the thermodynamic treatment. Physical theories of atomic structure and the energy relations of atoms and molecules are also included, though only to a very limited extent. The chemistry of strong electrolytes is developed on the basis of the theory of complete dissociation, while the section on weak electrolytes deals chiefly with protolysis and protolytic equilibria. In dealing with electrochemistry a rigorous treatment has been aimed at by using the concepts of electrochemical potential and the electromotive potential of metals. Finally, the book contains a short section on capillary

and colloid chemistry in which the behaviour of surfaces and colloidal solutions is dealt with, chiefly from a thermodynamic point of view.

While writing the book I have had the opportunity of discussing both scientific and educational points with Mr. E. Güntelberg, thus modifying the treatment in various respects. My thanks are also due to Mr. R. P. Bell for producing an excellent English translation of the book, and to the firm of William Heinemann for undertaking the publication of the English edition.

J. N. BRÖNSTED.

COPENHAGEN,
September, 1937

TABLE OF CONTENTS

	PAGE
CHAPTER I. THE LAWS OF THERMODYNAMICS . . .	1
1. THE FIRST LAW OF THERMODYNAMICS.	1
2. THE SECOND LAW OF THERMODYNAMICS	4
(a) Experimental Basis and Enunciation	4
(b) Reversible and Irreversible Processes	8
(c) The Transformation of Thermal Energy	11
(d) Temperature and Maximum Work	16
(e) Kinetic Considerations	17
3. THE THERMODYNAMIC FUNCTIONS	17
(a) Entropy	18
(b) Entropy and Equilibrium	23
(c) Other Thermodynamic Functions	23
(d) Open Systems. The Chemical Potential	26
(e) Applications	31
CHAPTER II. STATES OF AGGREGATION	36
1. THE GASEOUS STATE	36
(a) The Equation of State of Gases	36
(b) Internal Energy and Work	37
(c) The Gas Thermometer Scale and the Absolute Thermodynamic Scale	39
(d) The Specific Heat of Gases	39
(e) Adiabatic Changes	40
(f) Thermodynamic Functions for Gases	40
2. THE LIQUID STATE	42
(a) Surface Tension	42
(b) Viscosity	46
3. THE CRYSTALLINE STATE	47
4. CHANGES OF STATE. PHASE EQUILIBRIUM	48
(a) The Liquid-vapour Equilibrium	48
(b) Work and Heat of Vaporisation	48

	PAGE
(c) The Pressure-volume Diagram for the Process of Vaporisation	50
(d) The Vapour Pressure Curve	51
(e) The Equilibrium Crystal-liquid	53
(f) The Effect of Pressure on the Melting Point	55
(g) Allotropic and Polymorphic Transformations	55
(h) Application of Thermodynamic Functions to Phase Equilibria	57
(i) The Phase Diagram	58
 CHAPTER III. MOLECULAR THEORY	 61
1. THE BASIS AND AIMS OF MOLECULAR KINETIC THEORY	61
(a) The Avogadro Number, Atomic and Molecular Dimensions	61
(b) The Equipartition of Energy	62
2. THE KINETIC THEORY OF THE FLUID STATE	63
(a) Gas Pressure and the Kinetic Energy of Molecules	63
(b) The Distribution of Velocities and Energies. Maxwell's Law	66
(c) Boltzmann's Distribution Law	72
(d) The Specific Heat of Gases. The Quantum of Energy	74
(e) The Molecular Theory of the Liquid State	77
(f) The Transition from Liquid to Vapour	79
3. THE KINETIC THEORY OF THE CRYSTALLINE STATE	83
(a) Lattice Structure and Lattice Energy	83
(b) Specific Heat. Calculation from the Quantum Theory	84
 CHAPTER IV. MIXTURES	 89
1. PURE SUBSTANCES AND MIXTURES	89
2. GASEOUS MIXTURES	90
(a) Partial Pressures.	90
(b) Work of Mixing. The Thermodynamic Functions	91
3. LIQUID MIXTURES	94
(a) Work of Mixing	94
(b) General Relations between Differential and Integral Quantities	97

TABLE OF CONTENTS

xi

	PAGE
(c) The Chemical Potential. The Gibbs-Duhem Equation	100
(d) Ideal Mixtures	101
(e) Activities and Activity Coefficients	103
(f) Equilibrium between Liquid and Vapour	105
(g) Vaporisation and Distillation of Mixtures	107
(h) Vapour Pressure and Heat of Mixing	110
(i) Mutual Miscibility. Liquid-liquid Equilibrium.	111
(j) Ternary Systems	114
(k) The Solid-liquid Equilibrium	115
(l) The Vapour Pressure of Saturated Solutions	123
(m) Equilibrium between one Solid and two Liquid Phases	125
4. CRYSTALLINE MIXTURES	126
(a) The Vapour Pressures of Solid Solutions	127
(b) The Equilibrium Crystal Liquid	127
5. SYSTEMS OF IMMISCIBLE COMPONENTS	131
(a) Reactions between Solid Salts	131
(b) The Vapour Pressure of Salt Hydrates	131
6. THE PHASE RULE	137

CHAPTER V. THE THEORY OF DILUTE SOLUTIONS . 140

1. POTENTIALS AND ACTIVITIES	140
2. VAPOUR PRESSURE RELATIONS	144
(a) Henry's Law	144
(b) Raoult's Law	145
(c) Elevation of Boiling Point	147
(d) The Distribution Law	149
3. FREEZING POINT AND SOLUBILITY RELATIONS	149
(a) The Solubility of Solids	149
(b) Freezing Point Depression	151
4. OSMOTIC PRESSURE	153
(a) Definition and Measurement	153
(b) Osmotic Work	155
(c) Osmotic Pressure and Vapour Pressure	156
5. TERNARY MIXTURES	157

	PAGE
CHAPTER VI. HOMOGENEOUS EQUILIBRIA	159
1. BALANCED REACTIONS	159
2. THE LAW OF MASS ACTION	161
(a) Thermodynamic Derivation	161
(b) Kinetic Derivation	164
(c) Gaseous Systems	165
(d) Liquid Systems	168
(e) The Thermodynamic Law of Mass Action	170
3. HOMO-HETEROGENEOUS EQUILIBRIUM	170
(a) Thermodynamic Treatment	170
(b) The Equilibrium of Gaseous Systems with Solid Phases	171
(c) The Equilibrium of Liquid Systems with Solid Phases	172
4. THE EFFECT OF TEMPERATURE	174
(a) Thermodynamic Treatment	174
(b) Gaseous Systems	177
(c) Homo-heterogeneous Systems	182
CHAPTER VII. ELECTROLYTES	184
1. THE THEORY OF ELECTROLYTIC DISSOCIATION	184
2. STRONG ELECTROLYTES	186
(a) The Complete Dissociation of Strong Electrolytes	186
(b) Activities and Activity Coefficients	186
(c) The Variation of Activity Coefficients with Concentration	189
(d) The Solubility of Strong Electrolytes	194
(e) Determination of Activity Coefficients from Electromotive Force Measurements	198
(f) The Freezing Points of Salt Solutions	199
(g) Laws Governing Volume and Thermal Quantities	200
(h) The Theoretical Calculation of Activity Coefficients	201
3. WEAK ELECTROLYTES	202
(a) The Classical Theory	202
(b) The Acid-base Function. Protolysis	204
(c) Acidity and Basicity	208
(d) The Strength of Acids and Bases	212

TABLE OF CONTENTS

xiii

	PAGE
(e) Ampholytes	219
(f) Indicators	222
(g) Titration Theory	224
(h) Salt Effects	226
(i) The Solubility of Weak Electrolytes	228
4. THE EFFECT OF TEMPERATURE ON ELECTROLYTIC EQUILIBRIA	229
(a) The Solubility of Sparingly Soluble Salts	229
(b) The Dissociation of Weak Electrolytes	230
5. TERMINOLOGY OF PROTOLYTIC CONSTANTS	231
CHAPTER VIII. REACTION KINETICS	233
1. HOMOGENEOUS REACTIONS	233
(a) The Order of a Reaction	233
(b) Reversible Processes	236
(c) Unimolecular Reactions	237
(d) Bimolecular Reactions	239
(e) Ionic Reactions	241
(f) Homogeneous Catalysis	243
(g) Application of the Kinetic Theory of Gases	247
(h) The Effect of Temperature	252
(i) Chain Reactions	253
2. HETEROGENEOUS REACTIONS	255
(a) Diffusion Processes	255
(b) The Dissolution of Metals in Acids	256
(c) Heterogeneous Catalysis	256
3. PHOTOCHEMICAL REACTIONS	259
(a) Absorption of Light	260
(b) The Law of Photochemical Equivalence	260
(c) Photochemical Equilibrium	262
CHAPTER IX. ELECTROCHEMISTRY	264
1. ELECTROLYTIC CONDUCTION	264
(a) The Conditions for Electrolytic Conduction	264
(b) Electrochemical Reactions. Faraday's Law	265
(c) Electrolytic Transport	266
(d) Specific, Molar and Equivalent Conductivity	268

TABLE OF CONTENTS

	PAGE
(e) Strong Electrolytes	269
(f) Weak Electrolytes	273
(g) The Conductivity of Pure Water	274
(h) The Effect of the Solvent	274
2. ELECTRICAL POTENTIAL AND ELECTROMOTIVE FORCE	275
(a) Potential Differences between two Phases. The Electrochemical Potential.	275
(b) The Electromotive Potential of Metals	278
(c) Galvanic Cells	281
(d) Electromotive Reversibility	281
(e) The Thermodynamic Treatment of Galvanic Cells	283
(f) Concentration Cells	286
(g) Electrode Potentials	290
(h) Standard Electrode Potentials. The Electrochemical Series	293
(i) Electrodes of the Second Kind	296
(j) Oxidation-reduction Electrodes	298
(k) Gas Electrodes	302
(l) Liquid-liquid Junction Potentials	303
(m) Membrane Potentials and Electrolytic Membrane Equilibria	304
(n) Diffusion Potentials	310
(o) The Measurement of Single Potentials	314
(p) The Electrometric Determination of Hydrogen Ion Concentrations	315
(q) Polarisation and Over-voltage	318
CHAPTER X. AFFINITY	320
1. REACTION AFFINITY	320
(a) The Concept of Affinity	320
(b) The Principle of Affinity Determinations	322
(c) The Variation of Affinity with Temperature	324
(d) The Third Law of Thermodynamics	326
(e) Thermodynamic Functions and the Third Law	330
(f) The Calculation of Gas Equilibria from the Third Law	332
2. ATOMIC AND MOLECULAR PROCESSES	334
(a) Work of Ionisation and Electron Affinity	336
(b) Resonance Potentials	337
(c) Thermal Ionisation	338

TABLE OF CONTENTS

xv

	PAGE
(d) The Quantum Theory Model of the Atom	339
(e) The Relation between Radiation and Atomic Energies	341
(f) The Relation between Radiation and Molecular Energies	345
(g) Energy and Heat of Dissociation	346
 CHAPTER XI. SURFACE AND COLLOID CHEMISTRY	349
1. SURFACE CHEMISTRY	349
(a) The Adsorption of Gases	350
(b) Adsorption in Solution	354
(c) The Thermodynamics of Interfaces. Gibbs' Adsorption Equation	358
(d) Surface Films	366
2. COLLOID CHEMISTRY	367
(a) The Colloidal State	367
(b) The Formation of Colloidal Solutions	369
(c) The Thermodynamic Stability of Colloidal Solutions	371
(d) The Brownian Movement. The Diffusion of Colloids	374
(e) Sedimentation Equilibria. Determination of Avogadro's Number	377
(f) Sedimentation Velocity	379
(g) The Electric Charge of Colloids	379
(h) Osmotic Pressure	381
(i) Coagulation and Peptisation	385
(j) The Solubility Relations of Lyophilic Colloids	388

PHYSICAL CHEMISTRY

CHAPTER I

THE LAWS OF THERMODYNAMICS

1. THE FIRST LAW OF THERMODYNAMICS

THE first law of thermodynamics, or the *principle of conservation of energy* (J. R. Mayer, Joule, Colding), expresses the general experimental law that in a system which is completely isolated from its surroundings, the amount of energy remains constant for any kind of change taking place inside the system. As a consequence of this law, it is impossible to construct a "perpetual motion machine of the first kind," *i.e.*, a machine which produces energy from nothing.

The interconversion of the various forms of energy, on the other hand, takes place continually, and the treatment of such energy transformations constitutes the chief material of thermodynamics. The importance of such considerations in chemistry depends upon the close connection which exists between transformations of matter and transformations of energy.

The most important forms of energy dealt with in thermodynamics are mechanical, electrical, chemical and thermal energy. In addition to these, radiation energy is closely connected with chemical phenomena in certain cases. Mechanical energy can be further classified either as kinetic energy, *i.e.*, the energy which a body possesses in virtue of its motion (translation, rotation, etc.), or potential energy, *i.e.*, the energy possessed by a system of bodies in virtue of their position when there are forces acting between them. A corresponding distinction can be made in the case of electrical energy: thus two bodies which are charged to different electrical potentials constitute a system containing electrical potential energy, while an electric current may be said to contain electrical kinetic energy. Chemical energy (*e.g.*, the energy contained in a mixture of hydrogen and oxygen) may be regarded as potential in nature. As regards thermal energy, it will be shown later that it also can exist in forms differing essentially from one another.

The enunciation and treatment of the laws of thermodynamics

depends upon the measurement of amounts of energy. In the case of the forms of energy mentioned above, the amount of energy can be expressed as the product of two factors, an *intensity (or intensive) factor* and a *capacity (or extensive) factor*, and a quantity of energy can be measured by measuring each of these factors separately. Thus various forms of mechanical energy can be expressed as the product of force and distance, pressure and volume, or mechanical potential and mass, while thermal energy can be expressed as the product of temperature and entropy [3. a.], and electrical energy as the product of electrical potential and quantity of electricity. Chemical energy can similarly be split up into chemical potential and quantity of substance, but there is no general method of measuring the chemical potential in the same way that the intensity factors for the other forms of energy can be measured. The units of mechanical and electrical energy can be conveniently chosen as the product of the units of the appropriate intensity and capacity factors, but there is no corresponding unit for chemical energy, which can only be measured by converting it into other forms of energy.

Energy which can be determined in this way is also described as "work." In addition, there is thermal energy not possessing any potential character, which is termed "heat." This last form of energy is measured by the so-called "calorimetric" methods, which depend on the assumption that the amount of heat is not altered by passing from one body to another.

The most important energy units and their relation to one another are given in the following table :—

TABLE I. *Energy Units*

	Erg	g·cm	L. A.	g·cal	Joule
Erg	1	$1.020 \cdot 10^{-3}$	$9.87 \cdot 10^{-10}$	$2.389 \cdot 10^{-8}$	$0.9995 \cdot 10^{-7}$
g·cm	980.6	1	$9.678 \cdot 10^{-7}$	$2.342 \cdot 10^{-3}$	$9.801 \cdot 10^{-3}$
L. A.	$1.013 \cdot 10^8$	$1.0333 \cdot 10^8$	1	24.205	101.3
g·cal	$4.186 \cdot 10^7$	$4.269 \cdot 10^7$	0.04130	1	4.1842
Joule	$1.0005 \cdot 10^7$	$1.020 \cdot 10^7$	$9.87 \cdot 10^{-8}$	0.2390	1

The first column of the table contains the values of the various units in "absolute measure," i.e., expressed in the centimetre-gram-second (c.g.s.) system. The unit of energy in this system is the *erg*, defined as the product of one dyne, the unit of force, and 1 cm., the unit of length. A dyne is defined as that force which when acting on a mass of 1 gm. will produce in it an acceleration of 1 cm. per second per second.

The unit of heat given in the table is the 15° calorie, *i.e.*, the amount of energy necessary to heat 1 gm. of water from 15° C. to 16° C. This amount of energy is about 0.5% less than the zero calorie, which will heat 1 gm. of water from 0° C. to 1° C.

Chemical thermodynamics deals with systems which can undergo changes of a material or energetic nature. Such a system is termed a *thermodynamic system*. The real experimental proof of the first law of thermodynamics is based on a study of processes of such a kind that all the quantities of energy involved are susceptible of direct measurement, either by splitting them up into intensive and extensive factors, or by calorimetry. In such experiments it is found that the relation between the different kinds of energy is independent of the direction in which the transformation takes place, so that we can postulate that the sum of the amounts of work and heat which have disappeared is zero. Further, we attribute to every system taking part an *internal energy*, which can only be altered by interchange with the surroundings, and which is uniquely determined by the state of the system. If the sum of the amounts of work and heat which have disappeared differs from zero, we postulate a *change* in the internal energy of the system involved, which by definition fulfils the requirements of the first law: *i.e.*, for any alteration in the system the decrease in internal energy is put equal to the sum of the amounts of energy given up by the system to the surroundings. Thus if the heat absorbed by the system during the change is termed Q , the energy given up in other forms W ("work"), and the increase in the internal energy ΔE , then the first law of thermodynamics can be formulated as

$$\Delta E = Q - W \quad (1)$$

where it is of course essential that all the forms of energy taking part are expressed in the same units.

In general, when a system undergoes chemical or physical changes, at least two of the quantities in equation (1) will differ from zero. If one of the quantities is zero, equation (1) gives rise to three special cases:—

(1) $\Delta E = 0$. If a system passes through a *cycle*, *i.e.*, if after undergoing certain changes it finally returns to its initial state, or if the energy change involved in the processes performed is zero, then in all cases the heat absorbed must be equal to the work done, $W = Q$.

(2) $W = 0$. In most chemical processes, unless special arrangements are made, the work done will be zero, or at any rate very

small. The change in the internal energy is then measured by the heat absorbed, $\Delta E = Q$.

(3) $Q = 0$. If the system is in complete thermal isolation from its surroundings, any change taking place in the system is said to be adiabatic. We then have $\Delta E = -W$.

It should be mentioned that the internal energy E in equation (1) must not be considered as a special form of energy, but as composed of all the forms of energy already mentioned. It is not in general possible to state in what proportions the various forms of energy are present in a chemical system. In certain cases it is possible to demonstrate the presence of, *e.g.*, mechanical or electrical energy in the system, and a certain amount of thermal energy is always supposed to be present. However, the most important differences in the internal energy of a system depend on whether it can be characterised as potential or as non-potential energy. This question is more logically dealt with under the second law of thermodynamics.

2. THE SECOND LAW OF THERMODYNAMICS

(a) **Experimental Basis and Enunciation.** While the first law deals with the laws governing the transformations of different forms of energy which actually take place, the purpose of the second law (Carnot, Clausius, William Thomson) is to investigate the possibility of these transformations. As a result of these investigations it has been established that heat occupies a special position among the forms of energy, in that although all other energy forms can be converted directly to one another and directly to heat, the conversion of heat to other forms of energy (*e.g.*, to mechanical work) is subject to important restrictions.

These restrictions can be expressed in a general law which formulates the conditions under which heat can be transformed to work. This law is the second law of thermodynamics, and is usually stated by saying that it is impossible to construct a "perpetual motion machine of the second kind." This means that it is impossible to construct a machine which will convert "heat" to work without at the same time bringing about other changes of a material or energetic nature. On account of the large stores of heat in the earth, such a machine would in practice be just as useful producing work as a perpetual motion machine of the first kind.

We must now consider more closely this postulate and the experimental basis on which it rests. In order to understand the second law we must go more fully into the conception of energy as the product of an intensity factor and a capacity factor as mentioned in the last section. For the moment we will consider a material system which is completely cut off from its surroundings

with respect to the transport of both energy and matter. Such a system can in general be supposed to contain all the forms of energy mentioned above, but as long as each of the intensity factors (or *potentials*, as they may also be termed) has the same value throughout the system, the energy will have no tendency to become manifest, since there will be complete equilibrium preventing the occurrence of any energy changes. For example, if there are no differences of temperature or electrical potential, then no transformations of thermal or electrical energy will be able to take place. The occurrence of energetic processes thus demands the existence of differences of potential (or intensity factor) in the system in question, and the magnitude of the potential difference can be taken as a measure of the tendency to transformation of the appropriate form of energy. When such potential differences are present the system is said to contain *potential energy*. Similarly, the energy which can formally be said to be present at a uniform potential throughout the system will be termed *equipotential energy*.

In agreement with this idea, the energy changes which can take place in a system completely isolated from the surroundings consist of the migration of the capacity factors of energy from one potential to another. The spontaneous movement of the capacity factors is said to take place from a higher to a lower potential. Thus for example we may speak of the transference of entropy from one temperature to another, or of a quantity of matter from one chemical potential to another. Or the displacement of a mass from one mechanical potential to another (*e.g.*, the displacement of a weight in a direction at right angles to the earth's surface), or of a volume from one pressure to another. In the last case the spontaneous displacement takes place from a lower to a higher pressure, so that the pressure must be regarded as a negative potential. These processes cause changes in the amounts of energy in the forms in question, but provided that other forms of energy are excluded from taking part, the total amount of potential energy in the closed system will remain constant. Let us denote the mechanical, electrical, chemical, thermal and total potential energy by E_m , E_e , E_k , E_t and E_p respectively, the corresponding intensity factors or potentials by Φ , Ψ , μ and T , and the capacity factors by m , e , n and S . Let the change undergone by the system be infinitesimal, so that the state of the system shall not change during the process [3. a.]. Then if the indices 1 and 2 refer respectively to the *initial* and *final* states, we can write

$$\left. \begin{aligned} -dE_m &= (\Phi_1 - \Phi_2) dm, \\ -dE_e &= (\psi_1 - \psi_2) d\epsilon, \\ -dE_k &= (\mu_1 - \mu_2) dn, \\ -dE_t &= (T_1 - T_2) dS, \end{aligned} \right\} \quad (2)$$

while for the total potential energy E_p we have

$$dE_p = dE_m + dE_e + dE_k + dE_t, \quad (3)$$

and also

$$dE_p = 0. \quad (4)$$

It follows from the concept of potential energy involved in the first law that there is complete reciprocity with respect to the interconversion of the various forms of potential energy. Thus for example if E_e and E_k are constant, we have

$$dE_m + dE_t = 0$$

where each of the two terms may be either positive or negative. If dE_m is positive, the potential thermal energy which disappears will be transformed without loss to mechanical energy.

If we consider the postulate of the second law in the light of these considerations, it is clear that the "heat" which can only be transformed to work under certain conditions cannot be identified with potential thermal energy, which we have seen to be completely equivalent to the other forms of potential energy. The "heat" of the second law must thus be the thermal energy which we have termed equipotential, and which can also be described as "heat of constant temperature." It is in fact in general this kind of thermal energy which is intended when the term "heat" is used. Moreover, if this fact is recognised, it is clear that "heat" can no longer be said in the same way to occupy any special position among the forms of energy. For, as already noted, any form of equipotential energy (including mechanical, electrical or chemical energy) will be in the same state of impotence as regards the production of other forms of energy, in particular the production of mechanical work (*e.g.*, displacement of a weight against gravity) mentioned in enunciating the second law. If, therefore, the special position of "heat" is to be maintained, it must be sought elsewhere.

It must be noted here that the second law of thermodynamics does not state unconditionally that heat is unable to produce work (as might be expected from the equipotential nature of heat), but contains the important addition "without at the same time bringing about other changes of a material or energetic nature." However, we shall see later that the belief that heat can be converted to work

under certain conditions depends upon the fact that the usual formulation of the second law does not distinguish sufficiently between the potential and non-potential forms of thermal energy.

We must now investigate the possibility of converting heat into work when other processes are allowed to take part. In order to keep the various energy phenomena completely separate, we shall divide our experimental arrangement into a heat reservoir which can give up heat and which has the same temperature throughout, a work reservoir, in which work can be stored, and a reaction system placed between the two reservoirs. The boundaries of the reaction system are arranged so that any changes which might enable heat from the heat reservoir to be converted to work in the work reservoir take place within the system: *i.e.*, it is completely isolated from its surroundings except for its communication with the two reservoirs.

Let the amount of heat transferred from the heat reservoir to the reaction system be dq , the amount of work transferred from the reaction system to the work reservoir be dA , and the simultaneous increase in the total energy of the reaction system associated with any changes it undergoes be dE . We then have from the first law

$$dA = dq - dE.$$

If the experiment is carried out so that $dE = 0$ (which is of course always possible), then $dA = dq$, and the observed phenomena are thus the production of a certain quantity of work and the simultaneous disappearance of an equivalent amount of heat. However, these facts do not demonstrate that the work originates from the heat absorbed, since the reaction system has undergone a change during the process, and it would be just as feasible to assume that the work is produced from the total or "internal" energy of the system, the loss of internal energy being compensated for by the heat absorbed. Not only is it impossible on the present evidence to differentiate between these two possibilities, but there is no distinction between them from a purely logical point of view. This applies to the case $dE = 0$, but is obviously no less true when dE differs from zero.

However, the interpretation of the process described above becomes quite clear when we take into account the experimental fact that the amount of work obtained from a change taking place in a reaction system depends entirely on the state of the system and the nature of the change, and is not affected by the absorption of an equivalent or non-equivalent amount of heat by the system. Thus, for example, if the reaction system is a galvanic cell, then

the electrical energy produced by a definite infinitesimal change in the cell will be independent of the absorption of heat by the cell. Or, if the system contains a gas which expands into a vacuum of volume dv , giving work $p dv$, this work will be neither increased nor decreased by supplying a corresponding amount of heat to the system. The production of the work dA thus cannot be attributed to the heat dq transferred from the heat reservoir, but is entirely due to the change which takes place in the reaction system, where there is a decrease in internal potential energy equivalent to dA . The second law can therefore be enunciated in the following simplified form: *heat cannot be converted to work*. Or, if we choose to state that a perpetual motion machine of the second kind cannot be constructed, this can be defined simply as a machine or system *which produces work from heat*. At the same time the postulate can be *generalised*, since, as shown above, what we have said of heat applies to all equipotential forms of energy.

The law that heat cannot be converted to work may seem to be contradicted by the action of the so-called "heat engine," which is usually supposed to carry out such a conversion. However, we shall show in paragraph (c) that there is actually no contradiction.

The difference between heat and other forms of energy (which form the experimental basis for the second law) thus cannot be expressed by saying that it cannot be converted to work. This is also true of the other equipotential forms of energy, and such a conversion is equally impossible whether or not other material or energetic changes are allowed to take place. The peculiarity of heat among the other equipotential forms of energy depends on the fact that it can be *produced* from work. If the transformations in question are represented by the scheme

$$\text{Heat} \xrightleftharpoons[2]{1} \text{work},$$

then it is not the impossibility of process 1, but the possibility of process 2 which is specially characteristic of heat as a form of energy. The proof of this is connected with an important distinction observed in transformations of matter and energy, which is represented by the contrast between reversible and irreversible processes.

(b) Reversible and Irreversible Processes. To illustrate this contrast we must return to the closed system containing various forms of potential energy, described at the beginning of this section. If there are no differences of potential, all the amounts of potential energy can be taken as zero and the system will be in a state of complete equilibrium. This state might be called *neutral equilibrium*. If on the other hand there are potential differences and hence a

tendency to energy transformation in the system, then provided that forms of energy other than potential energy do not take part in the energy changes, these latter will be reciprocal, *i.e.*, the sum of all the amounts of energy produced and used up will be zero. This state can be termed a state of *balanced equilibrium*, since the various transformation tendencies balance one another in such a way that the total tendency is zero. Changes which take place in states of this kind are usually termed *reversible changes*. In order that such changes can take place in a state of balanced equilibrium it is obviously necessary that the transformation of the various forms of energy should be in some way *coupled together*, so that the transference of one capacity factor is necessarily associated with the simultaneous transference of another. This coupling together can in general only be brought about by special arrangements adapted to the particular transformations being dealt with, and it is of considerable importance for thermodynamical operations both in theory and in practice.

If on the other hand the various forms of potential energy are not coupled together in this way, then if the system contains potential energy there will be a positive transformation tendency, and transformation will take place *spontaneously*. The potential energy thus destroyed must by the first law give rise to the formation of equipotential energy. However, the capacity factors of all forms of energy other than heat (*e.g.*, amount of substance, quantity of electricity, mass) are all present in constant amount in a closed system, and equipotential energy of these parts therefore cannot be formed in the system. On the other hand, the *entropy*, which is the capacity factor of thermal energy, is not characterised by any such constancy, and the transformation of potential energy to equipotential energy must therefore result in the production of heat.

Hence, provided that no other potential energy is formed, the transformation of potential energy will result in the production of heat. *This expresses the special position of heat as a form of energy.* Since (as shown above) heat cannot be reconverted into potential energy, the conversion of the latter into heat is termed an *irreversible* process. The name *natural* process is also used in the same sense. In reality the heat is put completely out of action as a form of energy, since its conversion to potential energy is impossible. From this point of view the first law of thermodynamics takes on a somewhat different aspect.

We can summarise the result of the above considerations in the following scheme :—

$$\begin{array}{lcl}
 E_{p(1)} \xrightleftharpoons[2]{1} E_{p(2)}, & \left. \begin{array}{l} 1. \\ 2. \end{array} \right\} & \text{reversible,} \\
 E_p \xrightleftharpoons[2]{1} V, & \left. \begin{array}{l} 1. \\ 2. \end{array} \right\} & \begin{array}{l} \text{irreversible,} \\ \text{impossible.} \end{array}
 \end{array}$$

This expresses the conclusions arrived at above, viz., that inter-conversion of two forms of potential energy is always reversible, conversion of potential energy to heat irreversible, and conversion of heat to potential energy impossible.

The fact that irreversible processes are always accompanied by evolution of heat is exemplified by the heat produced by friction, and the heat evolved in a metal wire carrying an electric current. There are, however, numerous physico-chemical processes which appear to contradict this, since they are accompanied by a fall of temperature and are thus said to involve absorption of heat. For example, this is often the case when a salt is dissolved in water. It should however be noted that reversible processes in general may be accompanied by temperature changes, so that the latter are no criterion of the production or absorption of heat. In particular it must be recognised that a fall of temperature cannot be accepted as an indication of heat absorption: in fact, that an absorption of heat by a reacting system as a whole can never be the result of an energy process, since it would mean the conversion of equipotential energy to potential energy. What is actually observed, *e.g.*, in the case of a salt dissolving in water, is a fall of temperature which has the same origin as the fall of temperature observed when a liquid evaporates, *i.e.*, the conversion of one form of equipotential energy, which could be called "free" heat, to another form, which could be called "latent" heat. It is quite clear that heat absorption of this kind has no connection with any work which is simultaneously performed, since both positive and negative heat changes are met with in completely analogous processes of this kind. This distinction will be dealt with later from the point of view of kinetic theory.

The following scheme shows the forms of energy into which the total internal energy can be divided according to the above treatment:—

$$\text{Total } E \left\{ \begin{array}{l} \text{Potential } E \text{ (mechanical, electrical, chemical, thermal)} \\ \text{Equipotential } E \left\{ \begin{array}{l} \text{non-thermal } E \\ \text{heat} \left\{ \begin{array}{l} \text{free heat} \\ \text{latent heat} \end{array} \right. \end{array} \right. \end{array} \right.$$

Of the possible interchanges of these forms, it is the interconversion of mechanical, chemical and thermal potential energy and the

conversion of these forms of energy into heat which are of special significance for physico-chemical thermodynamics. The remainder of the present thermodynamical chapter will be devoted to considering interchanges of this kind.

(c) **The Transformation of Thermal Energy.** We shall consider the arrangement already mentioned in paragraph (a), in which an arbitrary material system is in balanced equilibrium with its surroundings so that the system can give up positive or negative work to a work reservoir and can receive positive or negative heat from a heat reservoir, while all other changes relevant to the arrangement take place inside the system. If the temperature of the system is constant throughout and hence identical with that of the heat reservoir, then for a reversible change in the system the work done upon the surroundings will be equal to the decrease in chemical and mechanical potential energy, while in an irreversible change a certain fraction of the potential energy will be converted to heat. A given isothermal process (i.e., an isothermal process with given initial and final states, for which all intermediate states have the same temperature as the initial and final states) will therefore perform its *maximum work* upon the surroundings when it takes place reversibly. Since this argument does not assume any particular mechanism for carrying out the process, the maximum work A done by an isothermal reversible process will be independent of the path by which the process takes place.

If potential thermal energy also takes part in the production of work, the temperature of the system cannot be the same throughout, since we have seen that potential thermal energy only exists in virtue of differences of temperature. We must thus assume that the system contains parts at different temperatures, most simply two heat reservoirs R_1 and R_2 with temperatures T_1 and T_2 , so that the work is produced by the transport of the capacity factor of thermal energy, *entropy*, from the reservoir with the higher temperature T_1 to the reservoir with the lower temperature T_2 . If this transport of entropy takes place reversibly, then the work A obtained by the transport of an amount of entropy S will be equal to the decrease in thermal potential energy, so that according to equation (2) we can write

$$A = (T_1 - T_2)S. \quad (5)$$

This is the maximum work associated with the transport of entropy, and by applying the first law to a reversible process it is seen that this maximum work must be independent of the path by which the transport takes place.

So far there is no essential difference between the calculations

of the work obtained from potential thermal energy and the work obtained from other forms of potential energy, since the decrease in the form of energy in question is quite generally equal to the product of a potential difference and the capacity factor transported. It would therefore be logical to treat all reversible processes under the first law of thermodynamics. However, the arrangement described in this paragraph is in principle identical with the so-called "heat engine," in which "heat" (*i.e.*, equipotential thermal energy) is said to be converted to work, and it is therefore of interest to continue with a more detailed analysis of the conditions under which the transport of entropy is converted to work.

We will for the moment consider the simplest form of the system, consisting only of the two heat reservoirs R_1 and R_2 at temperatures T_1 and T_2 . The system is in thermal communication with the heat reservoir R_0 which is at the temperature T_1 , and is also in communication with a work reservoir R_A .

In this system the process which produces work consists of a transport of entropy (amount S) from R_1 to R_2 , and according to our assumptions the work obtained is equal to the decrease in potential thermal energy associated with the transport of entropy, as given by equation (5).

If the quantities of heat lost by R_1 and given by R_2 are respectively q_1 and q_2 , then since thermal energy is expressible as the product of temperature and entropy, we have :—

$$S = \frac{q_1}{T_1} = \frac{q_2}{T_2} \quad (6)$$

so that the loss and gain of heat are in the same ratio as the temperatures of the corresponding reservoirs. Equations (5) and (6) give (in agreement with the first law),

$$A = q_1 - q_2. \quad (7)$$

If therefore we only wish to obtain a quantitative comparison of the quantities of energy concerned, we can consider the amount of heat q_1 given up by R_1 as being split up into two parts, of which one (q_2) is transferred from R_1 to R_2 , while the other ($q_1 - q_2$) disappears, giving rise to the formation of an equivalent amount of work, *i.e.*,

$$q_1(T_1) \begin{cases} \rightarrow q_2(T_2) \\ \rightarrow q_1 - q_2 = A \end{cases}$$

If we consider the heat is split up in this way, it might seem reasonable to say that the amount of heat $q_1 - q_2$ has been "converted"

to work. However, considerations of this kind give no information as to the actual origin of the work. In reality we can apply just the same treatment as was used in paragraph (c) in connection with the production of work by systems containing non-thermal potential energy, *i.e.*, the work can be considered as coming from a decrease of potential thermal energy *inside the system*, this potential thermal energy being itself formed by the transport of the amount of entropy S from T_1 to T_2 . The conversion of heat to work could be considered as established only if the absorption of an amount of heat $q_1 - q_2$ by the system (*i.e.*, R_1) were necessary to obtain the work A . Actually, the above considerations show that the production of work is independent of any such supply of heat from outside.

The heat engine thus constitutes no exception to the ordinary devices used for interconverting the different forms of energy. Its

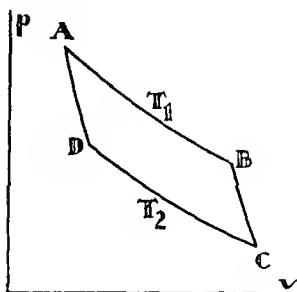


Fig 1

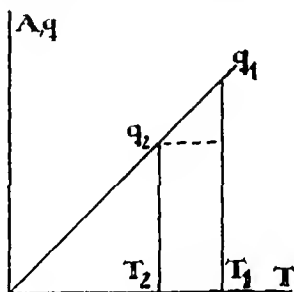


Fig. 2.

action in producing work is entirely analogous, *e.g.*, to that of an electric motor. In the first case the work is due to the descent of a quantity of entropy from a higher to a lower potential level, while in the second case it is due to the similar descent of a quantity of electricity.

As regards the detailed mechanism which can be used to obtain work by transport of entropy, there are various possibilities. The simplest would be to use an ideal thermo-couple: when the two junctions are placed in R_1 and R_2 this will produce electrical energy which can be directly converted to work in R_A , while the current will use an amount of heat q_1 in R_1 and give out amount of heat q_2 in R_2 . Or one can use the well-known *Carnot's cycle*, which gives a particularly clear picture of the conversion mechanism involved.

In Carnot's cycle a substance undergoes the reversible changes indicated by the four-sided figure $ABCD$ in the pressure-volume graph shown in Fig. 1. AB and CD are isotherms at temperatures

T_1 and T_2 , while BC and DA are adiabatic curves, i.e., curves representing changes which take place without absorption or emission of heat. If the heat absorptions along AB and DC are q_1 and q_2 respectively, then by the first law the work obtained in the whole cycle is

$$A = q_1 - q_2.$$

It is easily seen that the area of the figure $ABCD$ is a measure of this work. Further, we have from equation (6),

$$\frac{q_1}{q_2} = \frac{T_1}{T_2}. \quad (8)$$

Equation (6) or (8) can be used as a practical definition of temperature, since the temperature can be taken to be proportional to the heat absorption which takes place when an arbitrary system undergoes an isothermal and reversible change between two fixed adiabatic curves. The question of the universality of this definition is equivalent to the question of whether it is universally permissible to define thermal potential energy as the product of a quantity of entropy and the temperature. The latter question is formally answered in the affirmative by the consistency with which this definition of thermal potential energy can be applied in practice.

The general character of the equation is also shown by a closer examination of the Carnot cycle, which we consider applied to two independent systems between the same two temperatures T_1 and T_2 . Let the isothermal heat absorptions for the two systems be q_1 and q_1' at T_1 and q_2 and q_2' at T_2 . If the sizes of the systems are chosen so that $q_2 = q_2'$, then the heat absorptions q_1 and q_1' (and hence also the works obtained A and A') must be equal, since if this were not so, by letting the two processes oppose one another we could obtain a positive amount of work $\pm (A - A')$ at the expense of an amount of heat absorbed $\pm (q_1 - q_1')$, which would contradict the second law. Hence we must have

$$\frac{q_1}{q_2} = \frac{q_1'}{q_2'} \quad (9)$$

which shows the universal nature of the definition of temperature given above. It will be shown later that the so-called *thermodynamic temperature scale* defined in this way (Kelvin) is identical with the gas thermometer scale usually employed.

Fig. 2 shows graphically the content of equations (7) and (8). The figure also illustrates clearly the "efficiency" (defined as A/q_1) characteristic of a given "heat engine."

The efficiency is thus taken as the ratio of the work produced to the heat supplied. The reason for this terminology is that in practice the temperature T_1 is always higher than the normal temperature of the earth's surface, so that in order to maintain the temperature T_1 , heat must be supplied, e.g., by combustion. No corresponding arrangements are necessary in order to maintain the temperature T_2 , which is the temperature of the earth's surface. If the temperature of the earth's surface were used as the *higher* temperature in the heat engine, the technical problem would be the removal of q_2 and not the supply of q_1 , in which case there would hardly be the same grounds for considering heat as a form of energy which can be converted to other forms.

The above treatment deals with the energy changes associated with the transport of entropy from one temperature to another when the change is carried out reversibly. It will be seen that potential thermal energy is on the same footing as the other forms of potential energy, so that in general reversibility can be taken as typical of transformations in which interconversion takes place between different kinds of potential energy, but in which the total potential energy remains constant, i.e.,

$$dE_p = 0.$$

We shall now show that potential thermal energy is also on the same footing as the other forms of potential energy for changes which take place irreversibly by demonstrating that in all irreversible cases heat (i.e., equipotential thermal energy) is formed.

The transport of entropy can be caused to take place irreversibly simply by bringing the two heat reservoirs into thermal contact with one another when the process described as "transport of heat by conduction" takes place. Let us assume that the amount of heat thus transported from R_1 to R_2 is q_1 , so that this quantity of heat disappears from R_1 and appears in R_2 . If this process could be described merely as an irreversible transport of "heat," there would be no *production* of heat, as is the case for all other irreversible processes. It should however be noted that the transport of "heat," i.e., of energy, from one potential to another is a process exhibiting no analogy to the phenomena observed for other types of energy, since in the latter case it is the *capacity factor* of the type of energy in question which is transferred from one potential to another. If we maintain that this same process takes place for thermal energy, the process of "heat conduction" must be considered as a transference of entropy in both reversible and irreversible changes, and in the latter case heat must be formed at the expense of the potential thermal energy which is lost.

Let S be the amount of entropy, q_1 the heat lost by R_1 and q_2 the heat absorbed by R_2 , then

$$S = \frac{q_1}{T_1} = \frac{q_2}{T_2},$$

while the heat liberated when the process takes place irreversibly is identical with the work which would be obtained if the same amount of entropy were transported reversibly, i.e., it has the value $q_1 - q_2$. For a completely reversible change this heat must be liberated at the lowest temperature existing in the system, i.e., in R_2 . The total amount of heat absorbed by R_2 is hence $q_2 + (q_1 - q_2) = q_1$, i.e., the same as the heat lost by R_1 : this is of course required by the first law, and

corresponds to the simple conception of "heat conduction" as a direct transference of heat.

As regards the mechanism by which the amount of heat $q_1 - q_2$ is formed, we may assume that it takes place in the conductor connecting R_1 and R_2 and is produced by the "entropy current," just as heat is produced in a conductor carrying an electric current.

(d) **Temperature and Maximum Work.** If we use the Carnot cycle described above as a means of utilising potential thermal energy as work, then the quantity

$$A = S(T_1 - T_2)$$

can be termed the work done during the cycle. If the temperature difference is infinitesimal, this equation can be written

$$dA = SdT$$

or, introducing equation (6),

$$dA = q \cdot \frac{dT}{T}. \quad (10)$$

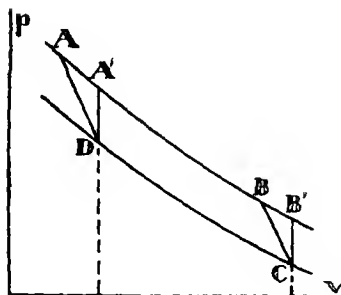


Fig. 3.

This amount of work can be represented by the area of the quadrilateral $ABCD$ (Fig. 3), where the isotherms AB and CD refer respectively to the temperatures $T + dT$ and T . It is easily seen that in the case considered here, where the two isotherms are only separated by an infinitesimal distance the area $ABCD$ can be put equal to the area $A'B'CD$, since the area of the triangles $AA'D$ and $BB'C$ are infinitesimal

compared with the area of the quadrilateral. Equation (10) will therefore hold for any arbitrary cycle between two isotherms with an infinitesimal temperature difference.

If the heating of the system between D and A' and between C and B' takes place at constant volume, then $dA = A_1 - A_2$, where A_1 is the work done along the isotherm $A'B'$, and A_2 is the work done along the isotherm DC . The work dA in equation (10) thus represents the increase in the work obtainable from a given process corresponding to an increase of temperature dT , the given process being the isothermal conversion of the initial state of the system corresponding to the point D , into the final state corresponding to the point C , the increase of temperature taking place

at constant volume. This result may be expressed by writing equation (10) in the form

$$q = T \left(\frac{dA}{dT} \right)_v. \quad (11)$$

As is shown by their derivation, q in the last two equations is the "reversible" heat absorption, or the "latent" heat of the process. If we introduce the expression for the first law, $\Delta E = q - A$, (11) becomes

$$A + \Delta E = T \left(\frac{dA}{dT} \right)_v, \quad (12)$$

where A is the maximum work. This expression summarises the two laws of thermodynamics.

(e) **Kinetic Considerations.** As will be shown in more detail in Chapter III, the internal energy of a chemical system in equilibrium consists partly of the kinetic energy of the thermal movement of the molecules, and in most cases partly of a kind of potential energy depending on the forces existing between the molecules or atoms. To this extent there is a formal analogy between internal thermal energy and the mechanical energy possessed by microscopic bodies in virtue of their position and motion. As shown above, the empirical difference between these forms of energy is shown by the impossibility of their reciprocal interconversion. From this point of view of molecular theory, this difference is due to the lack of order and direction which characterises the kinetic and potential components of thermal energy. The laws of thermodynamics are subject to the same qualifications as the impossibility of converting disordered molecular motion into ordered motion. It is clear that the possibility of such conversion is greater when only a few molecules are present than when (as in all practical cases) the number of molecules is very large: in fact, the problem itself vanishes when only one or two molecules are considered. This means that the second law of thermodynamics is really of a statistical nature, and can be formulated as a law of probabilities.

It should be carefully noted that what has been described as the potential energy of the molecules in an equilibrium system is essentially different from the potential energy described in section (1). "Molecular potential energy," as the first of these forms might be called, is not "potential" in the sense that work can be obtained from it. Its transformation to molecular motion gives rise macroscopically to "latent" heat.

3. THE THERMODYNAMIC FUNCTIONS

Among the thermodynamic quantities so far dealt with, the energy may be termed a *thermodynamic function*, since, as already stated, it is a quantity which depends only on the momentary state of the system in question. There exists a number of other

functions of which the same is true, which are useful quantities for deriving physico-chemical laws. The present section will be devoted to defining the most important of these functions and obtaining relations between them.

Since we are chiefly concerned with chemical applications, it will in general be permissible to exclude from consideration certain forms of energy, such as the mechanical energy which exists in virtue of gravitation, and the electrical energy which depends on the presence of free electric charges in the system. Similarly, we shall in general ignore capillary energy [II. 2. a.].

(a) **Entropy.** The concept of entropy has already been introduced in the preceding section as the capacity factor of thermal energy. This concept itself involves the assumption that entropy is a thermodynamic function, an assumption which is equivalent to the supposition that potential thermal energy is on the same footing as the other forms of potential energy. The character of entropy as a true thermodynamic function can be shown on a basis formally somewhat different by employing Carnot's cycle.

If a system is transformed reversibly from a state A of temperature T_1 to a state B of temperature T_2 by means of an isothermal and an adiabatic change, then the quantities of heat absorbed isothermally at T_1 and T_2 (q_1 and q_2 respectively) will, as shown above, satisfy the relation

$$\frac{q_1}{T_1} = \frac{q_2}{T_2}.$$

It will always be possible to build up the reversible transformation from state A to state B from a series of alternately isothermal and adiabatic processes, and if we apply the above expression to the zig-zag line thus obtained, we obtain

$$\sum_A^B \frac{q_1}{T_1} = \text{Constant}.$$

This means that the sum $\sum_A^B \frac{q_1}{T_1}$ will depend only upon the initial and final states A and B . Since any continuous curve connecting the points A and B can be considered as a zig-zag line of the kind described, we can also write for the integral along an arbitrary curve from A to B

$$\int_A^B \frac{dq}{T} = \text{Constant.} \quad (13)$$

The differential under the integral sign is thus a complete differential,

$$\frac{dq}{T} = dS, \quad (14)$$

and the quantity S thus introduced (first termed the entropy by Clausius) is a function only of the state of the system.

Equation (14) is subject to the important restriction that the absorption of the amount of heat dq takes place at the temperature T , so that it is not a sufficient condition for its validity that the initial and final temperatures of an infinitesimal change are both T . A process which conforms to the conditions of validity of (14) will be termed an *infinitesimal reversible process*, and in general we shall understand by an *infinitesimal process* one in which all the variables defining the state only undergo infinitesimal changes during the whole process.

Thus for a reversible change of a system from state A to state B , the increase in entropy is given by

$$\int_A^B \frac{dq}{T} = S_B - S_A, \quad (15)$$

where S_A and S_B are the entropies of the two states. In order for equation (15) to be valid, it does not matter whether the reversibility of the change from A to B is due to the presence of an internal equilibrium in the system (when no work will be performed by a change taking place at constant volume) or whether we are dealing with a balanced equilibrium in which potential energy is transformed to work by a displacement of the equilibrium.

If the latter is the case, *i.e.*, if the system is not in internal equilibrium and the change taking place can be made to produce a maximum amount of work by a suitable coupling between the forms of energy, then an incomplete utilisation of the available work will lead to an irreversible stabilisation of the system. Since the work done by the system in this case will be less than the work which is obtained when the same transformation of the system is carried out reversibly, on account of the law of conservation of energy the amount of heat supplied must also be less, while the

increase in both energy and entropy will be the same in both cases. We can therefore write for the two possible kinds of change

$$(\text{rev.}) \quad dq = TdS, \quad (16)$$

$$(\text{irrev.}) \quad dQ = TdS - \Delta, \quad (17)$$

where dq and dQ respectively are the amounts of heat supplied when the change takes place reversibly and irreversibly, and Δ is a quantity which is always positive.

We now introduce the first law, which for the two kinds of infinitesimal change of state assumes the forms

$$(\text{rev.}) \quad dE = dq - dA, \quad (18)$$

$$(\text{irrev.}) \quad dE = dQ - dW, \quad (19)$$

where dA and dW represent respectively the amounts of work actually obtained in reversible and irreversible transformations. Introducing these expressions into equations (16) and (17), we obtain

$$(\text{rev.}) \quad dE = TdS - dA, \quad (20)$$

$$(\text{irrev.}) \quad dE = TdS - dW - \Delta, \quad (21)$$

and hence

$$\Delta = dA - dW = dq - dQ. \quad (22)$$

Δ is thus the difference between the maximum work and the work actually obtained in carrying out the processes: similarly, from (16) and (17) it is the difference between the amounts of heat absorbed reversibly and irreversibly. If a given process is carried out in a series of different ways with gradually decreasing Δ , the course of the process will gradually approach reversibility. In other words, for a given process and a given quantity of substance, Δ can be taken as a measure of the irreversibility of the course of the process, since it represents the work lost (or the heat simultaneously produced) on account of the irreversibility.

It should be noted that the difference between reversible and irreversible processes expressed by the above equations only exists when the irreversibility is an *internal* irreversibility caused by differences of temperature, pressure or chemical potentials *in the system*. The equations for the thermodynamic functions will therefore be equally valid even if heat is supplied to the system from reservoirs having temperatures differing from the constant

temperature of the system, or if volume changes take place in contact with matter at pressures differing from the constant pressure of the system. Analogous statements are true for chemical changes in the system. This depends upon the fact that S , E , T and p are all true functions of the state of the system.

To illustrate further the rule that Δ has a positive sign, it is convenient to note that for systems of general chemical importance, irreversibility within the system can be due to the following causes:—

(1) Temperature differences in the system, which can bring about irreversible transport of heat. If the system is supposed to consist of two parts at different temperatures, and if entropy is transferred reversibly from the higher to the lower temperature without any simultaneous volume changes or chemical reactions, then as shown above the maximum work which can be obtained is

$$dA = (T_1 - T_2)dS,$$

while the work obtained by the irreversible transport of the same quantity of entropy is less than this.

(2) Pressure differences in the system, which can cause irreversible volume changes. It follows directly from the principle of irreversibility that such processes lead to a loss of available work.

(3) Differences in chemical potential in the system [3. d]. This can lead to irreversible chemical reactions or diffusion processes, which also take place with loss of available work.

If, therefore, only the above three causes of irreversibility are taken into account, the quantity Δ characteristic of irreversible processes can be written as the sum of three terms,

$$\Delta = \Delta(Q) + \Delta(v) + \Delta(n).$$

where Δ is positive, while the single terms composing Δ can be either positive or zero. Since equations (20) and (21) are valid for infinitesimal changes, the presence of the term TdS indicates that we have attributed to the system a temperature T which is uniform throughout. The term $\Delta(Q)$ thus disappears for changes of this kind. If the pressure is assumed constant throughout the whole system, the term $\Delta(v)$ also disappears. In this case—i.e., for irreversible processes taking place at constant temperature and pressure—we have the following equation analogous to (21),

$$dE = TdS - dW - \Delta. \quad (23).$$

In this equation Δ is defined as identical with $\Delta(n)$, i.e., as depending only on irreversible chemical processes or diffusion processes.

As already stated, Δ is a measure of the irreversibility of the path followed by the process, and will depend on the nature of this path. It is in principle possible to obtain the maximum work which can be performed by the chemical processes taking place in the system, by means of special methods (described later) depending on the possibility of coupling together the various forms of potential energy [2. b]. When therefore we speak of the reversible carrying out of a process which can also take place irreversibly, the use of such "special" methods is intended. If no such special arrangements are made, the only work performed will be that done *against* the pressure of the surroundings, and in this case we may say that the process is characterised by its maximum irreversibility. We then have

$$dW = pdv, \quad (24)$$

and hence

$$dE = TdS - pdv - \Delta_0, \quad (25)$$

where Δ_0 is the maximum value of Δ .

Δ_0 is a quantity which is completely fixed by the initial and final states of the system and is independent of the reaction mechanism, since the other quantities occurring in the equation are functions of the state of the system alone. Equation (25) is therefore always valid for a given change, independent of whether it takes place reversibly or irreversibly. It is only the *significance* of Δ_0 which is different in the two cases. If the change takes place reversibly Δ_0 represents the work obtained, while if it takes place completely irreversibly, Δ_0 is the work lost. In general we can characterise Δ_0 (defined by equation (25)) as a quantity which is positive for processes which *can* take place irreversibly (or, in other words, spontaneously), but whose magnitude is independent of the actual course taken by the process.

Equations (22) and (23) give

$$\Delta_0 = dA - pdv = dq - dQ_0, \quad (26)$$

where dQ_0 is the heat absorbed for maximum irreversibility, so that Δ_0 is the maximum work which can be obtained from an infinitesimal process taking place at constant volume.

If the process considered is not infinitesimal, but *finite*, then Δ_0 will not depend solely by the initial final states, but (as appears from equation (25)) will depend on the course by which the process takes place. The same is of course true of A and q .

(b) **Entropy and Equilibrium.** If equation (25) is applied to a system which is isolated from its surroundings, then $dE = 0$ and $dv = 0$. We can therefore write for a process taking place under these conditions (but otherwise unrestricted as to its path),

$$TdS = \Delta_0. \quad (27)$$

Natural or spontaneous processes (*i.e.*, processes which *can* take place irreversibly) will therefore obviously be accompanied by an increase of entropy. This is obviously true of both infinitesimal and finite processes. For processes which can only take place reversibly we have $\Delta_0 = 0$, and hence for processes of this kind

$$TdS = 0. \quad (28)$$

Equation (28) represents the condition that a process taking place inside a system can only take place reversibly. Since reversible processes always take place through a state of equilibrium, the condition for equilibrium can be expressed as

$$(dS)_{E,v} = 0. \quad (29)$$

This condition is generally valid, its special importance depending on the fact that it determines the equilibrium in a system which from a thermodynamic point of view is completely isolated from its surroundings.

(c) **Other Thermodynamic Functions.** According to equation (29), the conditions of equilibrium in a system can be found by investigating changes of entropy at constant volume and constant energy. In practice it is simpler to keep constant other parameters defining the state of the system, and it is possible to define corresponding thermodynamic functions which are more generally applicable in treating chemical equilibria.

The four following equations (of which only three are independent) serve to define such new thermodynamic functions in terms of energy and entropy.

$$\left. \begin{aligned} F &= E - TS, \\ G &= F + pv, \\ H &= G + TS, \\ E &= H - pv, \end{aligned} \right\} \quad (30)$$

All these functions were formulated by Willard Gibbs. The nomenclature used to describe them vary considerably. We shall employ the following nomenclature :—

F ; free energy,
 G ; thermodynamic potential,
 H ; heat content (enthalpy).

The term "free energy" was introduced by Helmholtz, but is sometimes used to describe the function G (G. N. Lewis).

The use of these functions is best seen by differentiating equation (30) and combining it with (25). This leads to the following set of fundamental equations (including (25) for the sake of convenience),

$$\left. \begin{aligned} dE &= TdS - pdv - \Delta_0, \\ dF &= -SdT - pdv - \Delta_0, \\ dG &= -SdT + vdp - \Delta_0, \\ dH &= TdS + vdp - \Delta_0, \end{aligned} \right\} \quad (31)$$

All these equations have of course the same range of validity as equation (25), and the meaning of Δ_0 is the same as was explained in the paragraph following equation (25).

It is seen from equation (31) that the increase in free energy (F) at constant temperature is equal to the work done reversibly on the system. Similarly, from equation (30) the increase in the heat content (H) at constant pressure is equal to the amount of heat absorbed, provided that only the amount of work pdv is performed. In the same way, the increase of energy at constant volume is equal to the amount of heat absorbed, provided that no work is performed, and the increase of energy at constant entropy is equal to the work done reversibly on the system. We can therefore write for a change of finite extent,

$$\left. \begin{aligned} (\Delta F)_T &= A, \\ (\Delta H)_p &= Q_0, \\ (\Delta E)_v &= Q_0, \\ (\Delta E)_S &= A, \end{aligned} \right\} \quad (32)$$

where Q_0 is the minimum absorption of heat corresponding to the maximum irreversibility, in agreement with (26).

As a further application of the fundamental equations (31) we shall investigate more closely the relation between the thermodynamic functions for changes of state in which $\Delta_0 = 0$, i.e., changes in which the maximum work of an infinitesimal process is expressed by pdv , and can thus be measured by a simple piston arrangement. As shown above, systems in which only this type of change can take place are in internal equilibrium. In agreement

with this condition we delete the term Δ_0 throughout in equation (31), giving

$$\left. \begin{aligned} dE &= TdS - pdv, \\ dF &= -SdT - pdv, \\ dG &= -SdT + vdp, \\ dH &= TdS + vdp, \end{aligned} \right\} \quad (33)$$

whence by partial differentiation,

$$\left. \begin{aligned} \left(\frac{\partial E}{\partial S}\right)_v &= T, & \left(\frac{\partial E}{\partial v}\right)_S &= -p, \\ \left(\frac{\partial F}{\partial T}\right)_v &= -S, & \left(\frac{\partial F}{\partial v}\right)_T &= -p, \\ \left(\frac{\partial G}{\partial T}\right)_p &= -S, & \left(\frac{\partial G}{\partial p}\right)_T &= v, \\ \left(\frac{\partial H}{\partial S}\right)_p &= T, & \left(\frac{\partial H}{\partial p}\right)_S &= v. \end{aligned} \right\} \quad (34)$$

These expressions will subsequently find extensive application in thermodynamical calculations.

The condition for equilibrium (29) was derived from (25), and it is clear that we can use (31) to derive other expressions for equilibrium conditions. Proceeding in the same way as in the deduction of (29), we obtain

$$\left. \begin{aligned} (dE)_{S,v} &= 0, \\ (dF)_{T,v} &= 0, \\ (dG)_{T,p} &= 0, \\ (dH)_{S,p} &= 0. \end{aligned} \right\} \quad (35)$$

We thus see that in a system kept at constant temperature and constant volume, irreversible processes will take place until the free energy has reached a minimum: similarly in a system kept at constant temperature and constant pressure, irreversible processes will take place until the thermodynamic potential has reached a minimum.

The equilibrium conditions involving E and H can be expressed in an analogous manner.

The fundamental equations (31) contain the term Δ_0 which is determined by the tendency to chemical reaction, and are therefore of decisive importance for a more detailed study of chemical equi-

libria, the quantity Δ_0 being related to the chemical potentials of the substances concerned. Such derivations are however most conveniently based on the laws valid for "open systems," which will be investigated in the next section.

By combining (34) with the definitions in equation (30) we obtain the two expressions

$$T \left(\frac{\partial F}{\partial T} \right)_v = F - E = -TS, \quad (36)$$

or

$$T \left(\frac{\partial G}{\partial T} \right)_p = G - H = -TS, \quad (37)$$

which can be re-written in the form

$$\left(\frac{\partial \left(\frac{F}{T} \right)}{\partial T} \right)_v = -\frac{E}{T^2}, \quad (38)$$

or

$$\left(\frac{\partial \left(\frac{G}{T} \right)}{\partial T} \right)_p = -\frac{H}{T^2} \quad (39)$$

The last two equations are in a convenient form for numerous applications.

By means of the relations given above and others which can be directly derived from them it is possible to calculate the changes in the thermodynamic functions, energy, entropy, free energy, etc., associated with a given change in the state of the system. No mention has been made of the absolute values of these functions, and such absolute values can in general only be determined by the conventional choice of a *standard state*, i.e., a completely defined state of the system in which definite arbitrary values are assigned to the thermodynamic functions. These values must of course be chosen so as to satisfy the fundamental equations (30).

(d) Open Systems. The Chemical Potential. For the changes of state so far considered it has been assumed that the system was "closed," in the sense that no matter was added or removed during the changes in question. Changes in the energy (and other thermodynamic functions) have been assumed to be caused solely by the transfer of heat (entropy) or work. The definition of energy is in reality based on this assumption. On the other hand, it is clear that the amount of substance is an important factor in determining

the energy and the other functions. Thus, for example, if the amount of substance is doubled, the internal (intensive) properties remaining unchanged, then the amount of energy will also be doubled. The question thus arises of how to formulate thermodynamic equations of general validity for *open systems*, i.e., systems which can gain or lose matter as well as heat and work.

The laws for the change of energy in an open system can be formulated by adding to the expression for dE in (33) general terms expressing the effect of amounts of substance added to the system. These terms depend both on the nature of the system and on the nature and quantity of the added substances. If we consider a homogeneous system composed of a series of components K_1, K_2, \dots , the effect of each of these components can be expressed as a product μdn , where dn is the amount of substance added and μ a factor determined by the nature of the system and of the substance. The part of (33) relating to change of energy thus becomes,

$$dE = TdS - pdv + \mu_1 dn_1 + \mu_2 dn_2 + \dots \quad (40)$$

The quantities μ_1, μ_2, \dots thus introduced are termed the *chemical potentials* of the substances in question. Their significance appears from the general relation by which the energy can be expressed in terms of the independent variables S, v, n_1, n_2 , according to the laws for partial differentiation, i.e.,

$$dE = \left(\frac{\partial E}{\partial S} \right)_{v,n} dS + \left(\frac{\partial E}{\partial v} \right)_{S,n} dv + \left(\frac{\partial E}{\partial n_1} \right)_{S,v,n} dn_1 + \left. \begin{aligned} & \left(\frac{\partial E}{\partial n_2} \right)_{S,v,n} dn_2 + \dots, \end{aligned} \right\} \quad (41)$$

where the index n means that the amounts of all the substances in the system are constant with the exception of the one which occurs in the denominator of the differential coefficient in question. Comparison of (40) and (41) gives

$$\left(\frac{\partial E}{\partial n_1} \right)_{S,v,n} = \mu_1, \quad \left(\frac{\partial E}{\partial n_2} \right)_{S,v,n} = \mu_2, \dots, \quad (42)$$

showing that the chemical potential of a component is defined as the increase in energy undergone by the system when unit amount of the component is added to the system in such a way that the entropy and volume of the system remain unchanged. It is also necessary that the unit amount shall be very small compared with the total mass of the system.

Instead of defining the potential by means of an energy change, as in equation (40), it can also be defined by its relation to other thermodynamic functions. Differentiation of (30) gives

$$\left. \begin{aligned} dF &= dE - TdS - SdT, \\ dG &= dF + pdv + vdp, \\ dH &= dG + TdS + SdT, \\ dE &= dH - pdv - vdp, \end{aligned} \right\} \quad (43)$$

and by combining this with (40) we obtain

$$\left. \begin{aligned} dF &= -SdT - pdv + \mu_1 dn_1 + \mu_2 dn_2 + \dots, \\ dG &= -SdT + vdp + \mu_1 dn_1 + \mu_2 dn_2 + \dots, \\ dH &= TdS + vdp + \mu_1 dn_1 + \mu_2 dn_2 + \dots \end{aligned} \right\} \quad (44)$$

It will be seen from (40) and (44) that we can write by analogy with (42)

$$\left. \begin{aligned} \mu_1 &= \left(\frac{\partial F}{\partial n_1} \right)_{T,p,n} = \left(\frac{\partial G}{\partial n_1} \right)_{T,p,n} = \left(\frac{\partial H}{\partial n_1} \right)_{S,p,n} = \\ &= -T \left(\frac{\partial S}{\partial n_1} \right)_{E,p,n} = -T \left(\frac{\partial S}{\partial n_1} \right)_{H,p,n}, \end{aligned} \right\} \quad (45)$$

thus obtaining a definition of μ in terms of F , G , H and S . These definitions characterise μ as an intensive quantity (*cf.* section 1). The first two are of special importance. We see that the potential of a component in a system can be expressed as the increase in free energy undergone by the system when unit amount of the component is added to the system, the temperature, volume and amounts of all the other components in the system remaining constant: or as the increase of thermodynamic potential undergone by the system when the same process is carried out at constant temperature, pressure and amounts of the other components. To determine μ completely it is therefore necessary to know the value of the function F (or G) in the standard state from which the transfer takes place. However, practical applications are in general concerned only with *differences* of potential, so that the standard state is usually without importance.

There is a particularly simple relation between μ and G . If the system only contains a single component, the expression for dG in (44) becomes

$$dG = -SdT + vdp + \mu dn, \quad (46)$$

and hence

$$\left(\frac{\partial G}{\partial n}\right)_{T,p} = \mu. \quad (47)$$

Further, for a chemical system of constant composition at constant T and p , G (like any other thermodynamic function) is proportional to the amount of substance, so that we can write

$$\left(\frac{\partial G}{\partial n}\right)_{T,p} = \frac{G}{n}. \quad (48)$$

Combining this with (47), we have

$$\mu = \frac{G}{n}, \quad (49)$$

i.e., the chemical potential of a pure substance is equal to the thermodynamic potential of the substance per unit amount.

If $\Sigma dn_i = 0$, while the individual changes dn_1, dn_2, \dots in the system differ from zero, then the transport of matter in the open system can be identical with a material reaction in the closed system. The energy changes in the system for these two kinds of process are given by (40) and (25) respectively as

$$dE = TdS - pdv + \mu_1 dn_1 + \mu_2 dn_2 + \dots,$$

and

$$dE = TdS - pdv - \Delta_0.$$

Since the two processes are identical, it follows that for a reversible or irreversible process taking place in the closed system

$$\Sigma \mu_i dn_i = -\Delta_0. \quad (50)$$

If the process consists solely on the transference of K_1 from a region where the potential is μ' , to a region where it is μ'' , then

$$(\mu' - \mu'')dn_1 = \Delta_0, \quad (51)$$

in agreement with the law that spontaneous transfer takes place from a higher to a lower potential. A similar formulation can be used if the process considered is of a more chemical nature.

If (20) and (40) are applied to a closed system, it follows that the maximum work is given by

$$dA = pdv - \sum \mu_1 dn_1, \quad (52)$$

or for constant volume by

$$dA = - \sum \mu_1 dn_1. \quad (53)$$

The same relations can be derived by applying the expression for dF in (44) to a closed system, since according to (32) $-dF$ is a measure of the maximum work.

In systems with uniform temperature and pressure, reversibility is always present for processes in which $\Delta_0 = 0$. Hence in addition to the equations already given, the equilibrium conditions for such a system can be expressed by

$$\sum \mu_1 dn_1 = 0, \quad (54)$$

where for each possible reaction the single amounts of substance dn_1, dn_2, \dots are connected by a stoichiometric equation [VI. 2. a.].

According to (54), the condition for equilibrium between two phases (*e.g.*, one liquid and the other vapour) is that for each single component contained in the system the chemical potential shall be the same in the two phases.

An important relation between the intensity factors T , p and μ can be obtained by considerations analogous to those on which equation (48) is based. At constant intensity factors, thermodynamic functions such as E , S , F and G are proportional to the amount of substance, and equations (40) and (44) can be directly integrated. We thus obtain for the function G the simple relation

$$G = \sum \mu_1 n_1. \quad (55)$$

If this equation is differentiated and the resulting expression for dG introduced in (44), we obtain the *fundamental equation* for the intensity factors of the system

$$SdT - vdp + \sum n_1 d\mu_1 = 0, \quad (56)$$

which is of great utility for thermodynamical calculations.

The use of the thermodynamic functions for calculating equilibria will be illustrated in the following sections. In order to express these equilibria in terms of directly measurable quantities such as temperature, pressure, volume, heat evolution, concentration, etc.,

it will be necessary to express the thermodynamic functions for the systems considered in terms of these variables. A complete description of the equilibrium is only possible when this can be done for all the variables concerned.

(e) **Applications.** This section will be devoted to the derivation of some formulæ of special interest for physico-chemical problems on the basis of the thermodynamic relations already given. It is necessary for this purpose to know the relation which exists between *changes of temperature, pressure and volume for a given system*, and we shall first derive the equation connecting these changes.

It is not possible to derive a general equation of state of the form

$$f(T, p, v) = 0.$$

However, if we consider, *e.g.*, a change of volume, the rule for partial differentiation gives the equation

$$dv = \left(\frac{\partial v}{\partial T}\right)_p dT + \left(\frac{\partial v}{\partial p}\right)_T dp, \quad (57)$$

or,

$$\frac{dv}{dT} = \left(\frac{\partial v}{\partial T}\right)_p + \left(\frac{\partial v}{\partial p}\right)_T \frac{dp}{dT}.$$

If this general equation is applied to changes at constant volume it is transformed to the equally general equation

$$\left(\frac{\partial v}{\partial T}\right)_p + \left(\frac{\partial v}{\partial p}\right)_T \left(\frac{\partial p}{\partial T}\right)_v = 0, \quad (58)$$

or,

$$\left(\frac{\partial v}{\partial p}\right)_T \left(\frac{\partial p}{\partial T}\right)_v \left(\frac{\partial T}{\partial v}\right)_p = -1.$$

This equation expresses the interdependence of the compressibility $\frac{1}{v} \left(\frac{\partial v}{\partial p}\right)_T$, the thermal pressure coefficient, $\left(\frac{\partial p}{\partial T}\right)_v$, and the coefficient of thermal expansion, $\frac{1}{v} \left(\frac{\partial v}{\partial T}\right)_p$.

We shall next investigate the concept of *specific heat*. This is defined as the ratio between the amount of heat absorbed by a substance and the consequent change in temperature. The specific heat can of course have different values, depending on the conditions under which the heat is supplied, and it is quite indeterminate as long as these conditions are not specified. Special importance attaches to the specific heats c_p and c_v , which are valid for supply of heat at constant pressure and constant volume respectively. It

is a necessary condition for the measurement of these quantities that no irreversible processes take place in the substance considered.

By combining the above definition with equation (32) we see that the specific heats can be expressed by

$$\left(\frac{\partial E}{\partial T}\right)_v = c_v, \quad (59)$$

and

$$\left(\frac{\partial H}{\partial T}\right)_p = c_p. \quad (60)$$

By combining these equations with (33) we obtain other expressions for the specific heats, *i.e.*,

$$T\left(\frac{\partial S}{\partial T}\right)_v = c_v, \quad (61)$$

$$T\left(\frac{\partial S}{\partial T}\right)_p = c_p. \quad (62)$$

Finally, combination of (33) with (61) and (62) gives

$$c_p = \left(\frac{\partial E}{\partial T}\right)_p + p\left(\frac{\partial v}{\partial T}\right)_p, \quad (63)$$

and

$$c_v = \left(\frac{\partial H}{\partial T}\right)_v - v\left(\frac{\partial p}{\partial T}\right)_v, \quad (64)$$

which relations can also be obtained directly from the definition of the specific heats, since at constant pressure we have

$$dq = dE + pdv$$

and at constant volume

$$dq = dH - vdp.$$

The difference between the two specific heats can be obtained in a simple form from the expression for a change of entropy,

$$dS = \left(\frac{\partial S}{\partial T}\right)_p dT + \left(\frac{\partial S}{\partial p}\right)_T dp, \quad (65)$$

which can also be written

$$\left(\frac{\partial S}{\partial T}\right)_v = \left(\frac{\partial S}{\partial T}\right)_p + \left(\frac{\partial S}{\partial p}\right)_T \left(\frac{\partial p}{\partial T}\right)_v. \quad (66)$$

Introducing (61) and (62) in this expression, we obtain

$$c_p - c_v = -T \left(\frac{\partial S}{\partial p} \right)_T \left(\frac{\partial p}{\partial T} \right)_v. \quad (67)$$

Applying equation (75) (derived below)

$$\left(\frac{\partial S}{\partial p} \right)_T = - \left(\frac{\partial v}{\partial T} \right)_p$$

we have for the difference between the specific heats in terms of the pressure coefficient and the coefficient of expansion,

$$c_p - c_v = T \left(\frac{\partial v}{\partial T} \right)_p \left(\frac{\partial p}{\partial T} \right)_v \quad (68)$$

Finally, with the help of (58) this can be re-written in the form,

$$c_p - c_v = T \frac{\left(\frac{\partial v}{\partial T} \right)_p^2}{\left(\frac{\partial v}{\partial p} \right)_T}, \quad (69)$$

which contains the compressibility in place of the less accessible pressure coefficient.

Equations (59) and (60) can be used for deriving the temperature coefficient of the heat evolved in transformations of chemical systems. If the initial and final states are denoted by the indices 1 and 2, we have

$$\left(\frac{\partial \Delta E}{\partial T} \right)_v = \left(\frac{\partial E_2}{\partial T} \right)_v - \left(\frac{\partial E_1}{\partial T} \right)_v = c_{v(2)} - c_{v(1)}, \quad (70)$$

and

$$\left(\frac{\partial \Delta H}{\partial T} \right)_p = \left(\frac{\partial \tilde{H}_2}{\partial T} \right)_p - \left(\frac{\partial H_1}{\partial T} \right)_p = c_{p(2)} - c_{p(1)}, \quad (71)$$

where ΔE and ΔH are respectively the increase in energy and heat content when the first system is transformed to the second at constant temperature. These quantities are identical with the heat absorbed during the process at constant volume and constant pressure respectively. These two equations express *Kirchhoff's Law*.

In order to derive the expressions for the variation of the entropy with pressure and volume (one of which has already been used) we can use the rule of differentiation according to which if a function of two independent variables is differentiated with respect to both variables, then the order of differentiation is immaterial. Thus, for example, if U is given by

$$dU = Xdx + Ydy, \quad (72)$$

we have

$$\left(\frac{\partial U}{\partial x}\right)_y = X, \quad \left(\frac{\partial U}{\partial y}\right)_x = Y, \quad (73)$$

and hence by the rule of differentiation given,

$$\frac{\partial^2 U}{\partial x \partial y} = \frac{\partial^2 U}{\partial y \partial x} = \left(\frac{\partial X}{\partial y}\right)_x = \left(\frac{\partial Y}{\partial x}\right)_y. \quad (74)$$

If we apply this result to the function G , given by (33) as

$$dG = -SdT + vdp,$$

we obtain

$$\left(\frac{\partial S}{\partial p}\right)_T = -\left(\frac{\partial v}{\partial T}\right)_p, \quad (75)$$

Similarly, using the function F and the variables T and v ,

$$\left(\frac{\partial S}{\partial v}\right)_T = \left(\frac{\partial p}{\partial T}\right)_v. \quad (76)$$

If (33) is written in the special form

$$\left(\frac{\partial E}{\partial v}\right)_T = T\left(\frac{\partial S}{\partial v}\right)_T - p, \quad (77)$$

insertion of (76) gives

$$p + \left(\frac{\partial E}{\partial v}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_v. \quad (78)$$

This important equation, which is valid for any homogeneous phase, is called the *thermodynamic equation of state*, and the quantity

$$\left(\frac{\partial E}{\partial v}\right)_T,$$

which obviously has the same dimensions as p , is called the internal pressure. The equation forms the basis for special equations of state, *e.g.*, for the real gases.

An equation containing -- v in place of p and H in place of E can be derived in an analogous way, giving

$$v - \left(\frac{\partial H}{\partial p}\right)_T = T \left(\frac{\partial v}{\partial T}\right)_p, \quad (79)$$

CHAPTER II

STATES OF AGGREGATION

1. THE GASEOUS STATE

(a) **The Equation of State of Gases.** The most obvious way in which gases differ from the other states of aggregation is in their property of decreasing indefinitely in density as the pressure approaches zero. The lower the pressure and the higher the temperature the more closely do gases approximate to the so-called ideal gas state, which is characterised by specially simple behaviour. In the case of some gases, *e.g.*, hydrogen and helium, the deviations from this state are negligible even under ordinary conditions. When we speak of the "gas laws," we mean the laws obeyed by gases in this ideal limiting state.

As the first of these laws we shall take *Boyle-Marriotte's law*, according to which the product of the pressure and the volume is constant for a given mass of gas at a given temperature. *The temperature on the gas thermometer scale* is defined in terms of this product by writing

$$pv = rT, \quad (1)$$

where r is a constant for the mass of gas considered. The size of the degree in this scale is determined by the convention that $T_1 - T_2 = 100^\circ$, where T_1 is the boiling point of water and T_2 its freezing point. Experiment has shown that the ratio between the corresponding products is

$$\frac{(pv)_1}{(pv)_2} = 1.366,$$

giving $T_1 = 373.1$, $T_2 = 273.1$.

For a given gas the constant r is of course proportional to the amount of substance present. If the amount of substance is one gram-molecule, experiment shows that for all ideal gases the constant has the same value R . We thus have for one gram-molecule of any ideal gas

$$pv = RT, \quad (2)$$

an equation which is termed the *equation of state* for ideal gases. R is called the *gas constant*. The molar volume (*i.e.*, the volume of

one gram-molecule) of an ideal gas at $T = 273.1$ and $p = 1$ atmosphere is found by experiment to be 22.41 litres, from which the value of the gas constant is calculated to be

$$R = 0.0821. \quad (3)$$

pv (which represents an amount of energy) is here reckoned in litre-atmospheres. If pv is in calories, we have

$$R = 1.985, \quad (4)$$

while if pv is in ergs,

$$R = 8.32 \times 10^7. \quad (5)$$

We have assumed here that the number of grams in a gram-molecule of gas has been fixed by chemical methods. It is possible to proceed differently, by *defining* one gram-molecule of gas as that amount which at very high dilution satisfies equation (2), using the numerical values given for R . This procedure is in some respects more satisfactory.

If the equation of state (2) is represented graphically in a p - v diagram, a rectangular hyperbola is obtained. Real gases will conform more closely to this equation the higher the temperature and the lower the pressure. At very low temperatures the equation ceases to be valid. Comparing different gases, the deviations will in general increase with increasing boiling point and increasing molecular complexity. If we introduce the "concentration" c as the number of gram-molecules present in unit volume, we have obviously $c = 1/v$, and equation (2) becomes

$$p = RTc. \quad (6)$$

(b) **Internal Energy and Work.** Experiments have shown that at constant temperature the energy content of an ideal gas is independent of the volume, *i.e.*,

$$\left(\frac{\partial E}{\partial v} \right)_T = 0. \quad (7)$$

This is shown by allowing the gas to pass spontaneously and adiabatically from one pressure to another: after equilibrium has been reached the temperature is found to be unchanged (Gay Lussac, Joule). If the state of the gas is changed from one isotherm to another, the energy change will therefore always be the same, independent of the way in which the change takes place. Further, if a change takes place along an isotherm, the work done will be equal to the heat absorbed,

$$q = pdv. \quad (8)$$

It follows from (2) and I. (30) that an equation analogous to (7) holds for the heat content H .

If the gas expands isothermally from volume v_1 to volume v_2 , the total work done will be

$$A = \int_{v_1}^{v_2} p dv. \quad (9)$$

If the amount of gas considered is 1 gram-molecule, we have by introducing (2)

$$A = RT \ln \frac{v_2}{v_1} = RT \ln \frac{p_1}{p_2}, \quad (10)$$

while the amount of heat simultaneously absorbed has the same value.

It is also of interest to consider the transfer of a finite amount of substance from a large reservoir A to another large reservoir B , and to investigate the work obtained when this transfer is carried out in different ways.

We will assume that the pressure in A is p and the pressure in B $p + dp$. We will further assume that any pressure changes produced in the reservoirs by the amounts of substance transferred are negligible compared with dp .

We shall first consider that the transfer is carried out in such a way that the volumes of A and B are kept constant. In the first stage the substance is removed from A by means of a cylinder with a movable piston, giving an amount of work $p v$. The pressure of the gas which has been removed is now altered to $p + dp$, giving an amount of work $p dv$, and it is finally introduced into B in the same way in which it was removed from A . The last operation gives an amount of work $-(p + dp)(v + dv)$, and we have in all

$$dA = p v + p dv - (p + dp)(v + dv)$$

or, since $dv dp$ can be omitted,

$$dA = -v dp. \quad (11)$$

If on the other hand the transfer is carried out in such a way that the pressures (and not the volumes) of A and B are kept constant, it is easily seen that the work obtained is simply

$$dA = p dv. \quad (12)$$

Both these formulæ are generally valid, since no equation of state was introduced in their derivation. If the gas law (2) is introduced, (11) and (12) become identical.

If we are dealing with ideal gases, the work done can also be calculated when there are finite pressure differences between the reservoirs. Integration of either (11) or (12) leads to the same

result: if the pressures in A and B are respectively p_1 and p_2 and the amount of substance transferred is 1 gram-molecule, then

$$A = RT \ln \frac{p_1}{p_2}. \quad (13)$$

(c) **The Gas Thermometer Scale and the Absolute Thermodynamic Scale.** According to (8) and (10) we have for the isothermal reversible volume change of a gas

$$q = A = RT \ln \frac{v_2}{v_1}. \quad (14)$$

Differentiation gives

$$dA = R dT \ln \frac{v_2}{v_1}, \quad (15)$$

and hence

$$\left(\frac{dA}{dT} \right)_v = \frac{q}{T}. \quad (16)$$

The temperature in equation (16) is measured by the gas thermometer. This equation is however identical with the expression previously obtained, I. (11), where T is measured by the absolute thermodynamic scale: hence the two scales must coincide. The size of the degree used in the thermodynamic scale is generally the same as that employed in the gas thermometer scale.

The same result is arrived at by introducing (7) into the thermodynamic equation of state, I. (78). This gives

$$\left(\frac{\partial p}{\partial T} \right)_v = \frac{p}{T},$$

which is identical with the relation between p and T derived from the equation of state $pv = RT$.

(d) **The Specific Heat of Gases.** For perfect gases we have from (7),

$$\left(\frac{\partial E}{\partial T} \right)_v = \left(\frac{\partial E}{\partial T} \right)_p. \quad (17)$$

These quantities are identical with the specific heat at constant volume.

For perfect gases the difference $c_p - c_v$ can be calculated from the general formula, I. (68)

$$c_p - c_v = T \left(\frac{\partial p}{\partial T} \right)_v \left(\frac{\partial v}{\partial T} \right)_p, \quad (18)$$

by introducing the equation of state (2). This gives for the specific heat per gram-molecule

$$c_p - c_v = p \left(\frac{\partial v}{\partial T} \right)_p = R. \quad (19)$$

If the specific heats are measured in calories, $c_p - c_v$ will therefore be 1.985.

The same result can be derived directly from (17), since this equation in conjunction with I. (18) shows that the difference

between the two specific heats is $p \left(\frac{\partial v}{\partial T} \right)_p$.

In equation (19), the difference on the left-hand side can be obtained directly from calorimetric measurements, while the term

$p \left(\frac{\partial v}{\partial T} \right)_p$ is given in mechanical units. This equation therefore

offers a possible method for determining the proportionality factor known as the mechanical equivalent of heat (J. R. Mayer).

Important relations concerning the specific heats of perfect gases are obtained from kinetic theory [III. 2. d.].

(e) **Adiabatic Changes.** For a reversible adiabatic change $dS = 0$, and hence from I. (33) :

$$dE + p dv = 0.$$

Introducing $dE = c_v dT$, we have

$$c_v dT + p dv = 0. \quad (20)$$

For a perfect gas we can introduce the equation of state (2), giving

$$c_v d \ln T + R d \ln v = 0.$$

Introducing (19) and assuming that $\frac{c_p}{c_v} = k$ we have

$$T v^{k-1} = \text{Constant}, \quad (21)$$

$$\text{giving on introducing (2), } p v^k = \text{Constant}, \quad (22)$$

or

$$T^k p^{1-k} = \text{Constant}. \quad (23)$$

(f) **Thermodynamic Functions for Gases.** For an arbitrary change in the state of a perfect gas, I. (59) gives

$$dE = c_v dT. \quad (24)$$

The entropy change can be calculated from I. (33) :

$$dE = T ds - p dv,$$

by means of (24), the equation of state (2) and the relation (19) between the specific heats. Remembering that $d\ln T = d\ln v + d\ln p$, this gives

$$\left. \begin{aligned} dS &= c_v d\ln T + R d\ln v, \\ dS &= c_p d\ln T - R d\ln p, \\ dS &= c_v d\ln p + c_p d\ln v. \end{aligned} \right\} \quad (25)$$

Except at low temperatures, the specific heat c_v can be considered as approximately independent of temperature. This enables us to obtain approximately valid expressions for energy and entropy by integrating (24) and (25),

$$E = c_v T + I_1, \quad (26)$$

$$S = c_v \ln T + R \ln v + I_2. \quad (27)$$

(27) and I. (30) give for the other thermodynamic functions

$$F = -c_v T \ln T + (c_v - R \ln v - I_2) T + I_1, \quad (28)$$

$$G = -c_v T \ln T + (c_p - R \ln v - I_2) T + I_1, \quad (29)$$

$$H = c_v T + I_1. \quad (30)$$

An expression for the chemical potential of a gas can be derived from I. (34),

$$\left(\frac{\partial G}{\partial p} \right)_T = v,$$

by introducing I. (49) and the equation of state (2). This gives for 1 gram-molecule

$$\left(\frac{\partial \mu}{\partial p} \right)_T = \left(\frac{RT}{p} \right),$$

or,

$$d\mu = RT d\ln p, \quad (31)$$

valid for constant temperature. Integration gives

$$\mu = RT \ln p + i, \quad (32)$$

where i is a constant depending on the temperature but not on the pressure. The constant is obviously equal to the value of the potential at the temperature T and at unit pressure. If i represents a constant independent of both T and p , then (32) is only valid at constant temperature. The relation can also be derived directly from (29).

If (32) is applied to two states of 1 gram-molecule of gas at the same temperature, i can be eliminated, giving

$$\mu_1 - \mu_2 = RT \ln \frac{p_1}{p_2}. \quad (33)$$

2. THE LIQUID STATE

In contrast to gases, liquids are characterised by forces of cohesion depending upon intermolecular attractions, in virtue of which they can in general exist with finite volume not only at zero pressure but also at negative pressures. The gaseous and liquid states cannot however be said to differ in principle, and it is possible to carry out continuous transitions from one to the other. Cohesion causes liquids to have a relatively high density and changes in the external pressure or the temperature only produce a small volume change for the same reason.

(a) **Surface Tension.** The presence of cohesion is particularly noticeable at the surface of a liquid, since here the action of the cohesive forces is more one-sided than in the interior of a mass of liquid. The surface tension γ is defined by considering a unit of length in the surface of the liquid: the surface tension is then the force acting at right angles to the unit of length and parallel to the surface. The surface tension thus has the dimensions of force/length. On multiplying it by the area of the surface a quantity having the dimensions of energy is obtained.

It is convenient to note here the different forms of mechanical energy which are of importance for physico-chemical phenomena, namely,

- (1) $k \cdot l$ force \times distance
- (2) $\frac{k}{l} \cdot l^2$ tension \times area
- (3) $\frac{k}{l^2} \cdot l^3$ pressure \times volume.

It is the second of these forms which is directly concerned in surface phenomena.

The surface can constitute the interface between the liquid considered on the one hand, and on the other hand a vacuum, the vapour of the liquid, the atmosphere, or other gases, liquids or solids. It is to be expected that γ will be different in these different cases. The surface tension is thus really a property of two phases, and is more correctly described as an interfacial tension. However, if one of the phases is gaseous it will in general have practically no effect on the value of the tension.

Surface tension can be determined by measuring directly the tension in these films of liquid, or by the rise of the liquid in capillary tubes. In the latter case the equation

$$2\pi r\gamma = \pi^2 h\rho$$

(where r is the radius of the tube and ρ the density of the liquid) expresses the fact that the surface force is equal to the weight of liquid supported. It gives for the surface tension

$$\gamma = \frac{1}{2}rh\rho \quad (34)$$

so that for a given liquid the rise is inversely proportional to the radius.

As an example we shall give the values of the surface tension in ergs per sq. cm. for a number of liquids at 20°.

TABLE I. *Surface Tension*

Liquid.	γ .
Water	72.8
Mercury	440
Alcohol	22.0
Glycerol	64
Benzene	28.9
Ether	16.9
Acetic acid	23.5

The surface tension between a liquid and its vapour will always decrease with rise of temperature and will become zero at the critical point.

For "normal" liquids [IV. 3. i.] the following relation has been found

$$\gamma v^{\frac{2}{3}} = k(T_0 - T), \quad (35)$$

where v is the molar volume, T the experimental temperature, T_0 the critical temperature and k a constant independent of the nature of the liquid. Since $v^{\frac{2}{3}}$ is proportional to the surface of, *e.g.*, a sphere containing one gram-molecule, the above equation states that the temperature coefficient of the molar surface work has a constant value for "normal" liquids. Equation (35) is however only very approximate, and is inapplicable to liquids like water and alcohol which deviate largely from "normality."

A quantity which experience shows to be independent of temperature is

$$\frac{M}{D - d} \gamma^{\frac{1}{2}}.$$

where M is the molecular weight, D the density of the liquid and d that of the vapour. This quantity is called the "parachor," and is found to be an additive function of constants characteristic of the single atoms and groups of atoms of which the molecule is built up.

The increase of the surface of a liquid requires an amount of work which is equal to the product of the surface tension and the increase of area. The process will also be accompanied by thermal phenomena. The energy of a given quantity of a liquid will therefore be dependent upon the extent of its surface. In the thermodynamic formulæ developed so far, we have used the expression $dA = pdv$, and in general taken no account of the surface as possessing any thermodynamic properties. However, the energy changes associated with changes of surface are in general negligible compared with the other amounts of energy concerned in the equations. Thus the specific surface work for water (*i.e.*, the work necessary to increase the surface of water by 1 sq. cm. at constant volume) is only 72 ergs, or about 1.7×10^{-6} calories. On the other hand, for very large surfaces, such as occur in emulsions and colloidal solutions, the surface energy becomes of importance.

If the surface work is taken into account, we must write in place of $dA = pdv$,

$$dA = pdv - \gamma d\sigma,$$

where σ is the area of the surface. If the change of surface takes place at constant volume, we therefore have

$$(dA)_v = -\gamma d\sigma. \quad (36)$$

All the previously derived thermodynamic equations can therefore be transformed into new equations by writing throughout σ instead of v , and $-\gamma$ instead of p , provided that the processes considered take place at constant volume. Thus the thermodynamic equation of state (I. 78)

$$\left(\frac{\partial E}{\partial v}\right)_T + p = T\left(\frac{\partial p}{\partial T}\right)_v,$$

must be written as

$$\left(\frac{\partial E}{\partial v}\right)_{T,\sigma} + p = T\left(\frac{\partial p}{\partial T}\right)_{v,\sigma}, \quad (37)$$

when the effect of the surface is taken into account, and this leads to

$$\left(\frac{\partial E}{\partial \sigma}\right)_{T,v} - \gamma = -T\left(\frac{\partial \gamma}{\partial T}\right)_{v,\sigma}. \quad (38)$$

This equation relates the surface tension and its temperature coefficient to the quantity

$$E_\gamma = \left(\frac{\partial E}{\partial \sigma}\right)_{T,v},$$

which is the specific surface energy. Since γ decreases with increasing temperature for the liquid-vapour interface, E_γ will be greater than γ : i.e., the increase in energy is greater than the work done. The difference between these two quantities is stored in the system as "latent" heat, analogous to latent heat of evaporation.

According to (35), γ is approximately a linear function of the temperature; hence from (38) the specific surface energy is approximately independent of temperature.

It is a matter of experience that small drops unite spontaneously to form larger ones. This phenomenon depends upon the fact that the union of the small drops is accompanied by a decrease of surface, so that surface work can be obtained from the process. The effect of the radius of the drops can be shown in the following way.

From the formulae $\sigma = 4\pi r^2$, $v = \frac{4}{3}\pi r^3$, we see that the radius r , the surface σ and the volume v of a sphere satisfy the equations

$$\frac{1}{3} \frac{dv}{v} = \frac{1}{2} \frac{d\sigma}{\sigma} = \frac{dr}{r},$$

and

$$\frac{v}{\sigma} = \frac{r}{3}.$$

This gives

$$d\sigma = \frac{2}{r} dv,$$

and for the change of surface when a positive volume v is transferred from the first sphere to the second,

$$d\sigma = d\sigma_2 - d\sigma_1 = 2 \left(\frac{1}{r_2} - \frac{1}{r_1} \right) dv,$$

This shows that the surface is decreased when $r_2 > r_1$, i.e., when the transference takes place from a smaller to a larger drop. The numerical value for the work of transfer is obtained by multiplying $d\sigma$ by γ . If $r_2 > r_1$ the total work for the union of the small drop with the large one becomes

$$\gamma d\sigma_1 = 4\pi r_1^2 \gamma. \quad (39)$$

The work of transfer $-\gamma d\sigma$ can also be expressed by means of equation (52), so that we can write for one gram molecule

$$-\frac{v}{dv} \gamma d\sigma = RT \ln \frac{p_1}{p_2},$$

whence

$$\ln \frac{p_1}{p_2} = -\frac{2v\gamma}{RT} \left(\frac{1}{r_2} - \frac{1}{r_1} \right). \quad (40)$$

v is here the molar volume of the liquid, r_1 , p_1 and r_2 , p_2 the radii and vapour pressures of the two drops. It will be seen that the smaller drop has the greater vapour pressure, and that the ratio of vapour pressures can be calculated if the molar volume and surface tension of the liquid are known.

Two drops of equal size will not have any tendency to unite, since they have the same vapour pressure. If however the process is started by the passage of a small quantity of substance from one drop to the other, it will continue with ever-increasing tendency until the two drops have become one.

(b) **Viscosity.** The extent to which a liquid is mobile is expressed by the term viscosity, a high viscosity denoting a great resistance to change of form. For some substances (*e.g.*, glass) the viscosity is so high that the mechanical properties of the substance approach those of the crystalline state. Substances of this kind are termed *amorphous*.

If the space between two plane parallel plates distant l apart is filled with a liquid, and it is necessary to apply to one of the plates a force k to produce a velocity h relative to the other plate, then for small values of l the following equation will hold,

$$h = k \frac{l}{\eta}, \quad (41)$$

where η is the coefficient of viscosity of the liquid. The coefficient of viscosity is thus defined as the force per unit area necessary to produce unit relative velocity when the plates are unit distance apart.

The viscosity can be determined from the rate at which the liquid flows through a narrow tube, or from the rate at which a sphere sinks in the liquid. In the latter case the motion obeys *Stokes' law* (XI. 2. d.),

$$B = \frac{1}{6\pi\eta r}, \quad (42)$$

$B = h/k$ is the "mobility," *i.e.*, the velocity produced by unit force, and r is the radius of the sphere. The application of this formula to particles of various kinds is of considerable importance in several branches of physical chemistry. The following table gives the viscosity of a few liquids at about 20° , expressed in abso-

lute units $\left(\eta = \frac{g}{\text{cm} \cdot \text{sec}}\right)$.

TABLE 2. *Viscosity*

Liquid.						
Ethyl alcohol	0.012
Ethyl ether	0.0022
Water	0.010
Glycerine (86%)	10.7
Mercury	0.016

3. THE CRYSTALLINE STATE

Crystals are characterised by a fixed geometrical form, which is not only apparent externally but is a property belonging to the crystal as a whole: this is shown, for example, by the existence of definite planes of cleavage. They are also characterised by a certain elasticity, *i.e.*, if they are deformed (within certain limits) they tend to return to their original form when the deforming force is removed. We can also say that the properties of crystals are in general dependent on the direction relative to the geometrical axes characterising the crystalline form: this is true, *e.g.*, of the coefficient of expansion and the compressibility. Such properties as the tendency to evaporate or dissolve are also different for different directions within the crystal. For this reason crystals are described as *anisotropic* substances, in contrast to the two other states of aggregation, gaseous and liquid, which together constitute the *fluid* or *isotropic* state. This difference depends upon the difference in the inner structures of isotropic and anisotropic substances, and can be accounted for on the basis of molecular theory.

While gases approach the perfect gas state with increasing temperature and decreasing pressure, crystals possess the simplest properties at low temperatures and high pressures. Under these conditions they approach the so-called *ideal crystal state*, in which the mechanical and thermal properties are constants independent of pressure and temperature. Thus the coefficient of expansion and the compressibility (which even under ordinary experimental conditions are very small compared with the corresponding quantities for gases) will approach zero with decreasing temperature. It is however difficult in practice to reach the ideal crystal state.

According to the simple rule of Dulong and Petit, the atomic heat (*i.e.*, the product of the specific heat and the atomic weight) of an element in the crystalline state should be about 6.2 at ordinary temperatures. For crystalline compounds the rule of Kopp and Neumann states that the molecular heat is approximately an additive function of the atomic heats. These rules are not however universally applicable, and in some cases, *e.g.*, for carbon and boron, there are considerable deviations.

At low temperatures the specific heat of crystalline substances falls asymptotically to zero, the curves for a few substances being

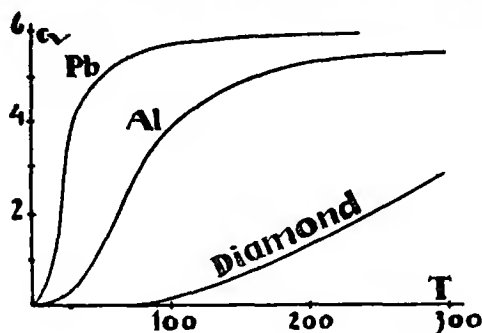


Fig. 1.

shown in Fig. 1. At very low temperatures the specific heat can be expressed by the equation

$$c_v = \alpha T^3, \quad (43)$$

where α is a constant characteristic of the substance. The temperature dependence of the specific heats of solids has been of great importance for the molecular theory of this state of aggregation, and will be considered further in Chapter III.

4. CHANGES OF STATE. PHASE EQUILIBRIUM

(a) **The Liquid-vapour Equilibrium.** If a substance is present as two phases simultaneously, *e.g.*, as liquid and as vapour, then if the temperature is fixed there can only be equilibrium between the phases if the pressure has a certain definite value called the vapour pressure. If the pressure is greater or less than this, condensation or vaporisation takes place respectively until the equilibrium pressure is reached. Vaporisation is accompanied by absorption of heat, *i.e.*, if the temperature is to be kept constant during the process, heat must be supplied. The vapour pressure increases with increasing temperature, and the curve representing this increase will be investigated below.

(b) **Work and Heat of Vaporisation.** If a liquid is vaporised in a vacuum, no work is done. If it is vaporised reversibly at an external pressure equal to the vapour pressure p , 1 gram-molecule may be said to do work $A = p(v - v_1)$, where v is the volume of the vapour and v_1 that of the liquid. This expression can be simplified to

$$A = pv = RT, \quad (44)$$

if v_1 can be neglected in comparison to v . This work is thus independent of the magnitude of the vapour pressure, and hence of the nature of the liquid, depending only on the temperature.

The heat of vaporisation q is the heat absorbed during isothermal vaporisation when there is equilibrium between the two phases. In contrast to the work of vaporisation, it depends on the nature of the liquid. If the decrease in energy on vaporisation is termed $-\Delta E$, we have the equation

$$q = A + \Delta E,$$

and the greatest heat of vaporisation is thus obtained for reversible vaporisation. Under ordinary conditions (*i.e.*, when the liquid is in contact with saturated vapour) the process of vaporisation takes place reversibly, and the heat of vaporisation determined calorimetrically is therefore equal to the reversible heat of vaporisation. This is also the case when the vaporisation takes place in contact with the atmosphere and is accompanied by irreversible diffusion of vapour, since the work done is RT in this case also. The work of vaporisation is as a rule only a small fraction of the heat of vaporisation.

When a vapour is condensed to liquid, an amount of heat is evolved equal in magnitude to the heat of vaporisation. This "heat evolution" does not originate in the same way as the heat produced by the irreversible transformation of other forms of energy: this is shown (for example) by its symmetrical relation to the heat of vaporisation, which is a negative heat evolution. Both these heat phenomena (as well as analogous phenomena associated with other changes of state and similar processes) are best characterised as the extensive thermal quantities which must be supplied or removed in order that the process shall be able to take place isothermally. It would actually be in many respects advantageous to deal with the entropy change associated with vaporisation (and other changes of state), rather than with the corresponding amount of heat. This is shown, for example, by the fact that in the equations governing such changes the amount of heat appears divided by T .

The variation of the heat of vaporisation with temperature is calculated in the following way. By differentiating the last equation we find for the reversible heat of vaporisation

$$\frac{dq}{dT} = \frac{dA}{dT} + \frac{d\Delta E}{dT}.$$

We also have from (19) and (44)

$$\frac{dA}{dT} = R = c_p - c_v,$$

and from I. (70)

$$-\frac{d\Delta E}{dT} = c_1 - c_v, \quad (45)$$

where c_1 , c_p and c_v are the specific heats of the liquid, the vapour at constant pressure and the vapour at constant volume, respectively. These equations give

$$\frac{dq}{dT} = c_p - c_1. \quad (46)$$

The heat of vaporisation will thus decrease with increasing temperature (since the specific heat of a liquid is greater than that of the corresponding vapour), and will approach zero as the temperature approaches the critical point.

(c) **The Pressure-volume Diagram for the Process of Vaporisation.** In the p - v diagram in Fig. 2 the curves $ABCD$ and $A'B'C'D'$ represent two isotherms for the fluid state. AB is the isotherm for the

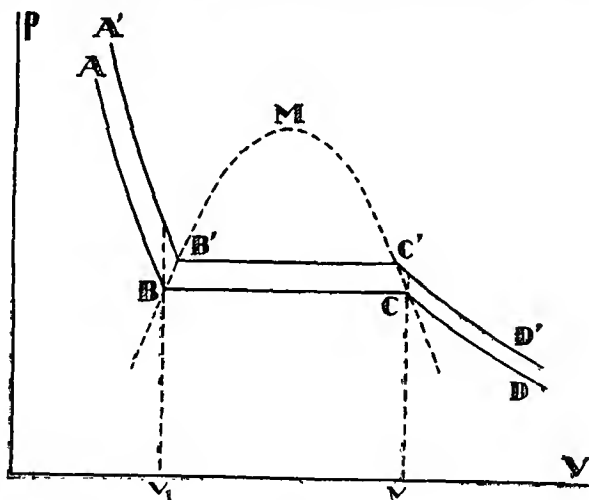


Fig 2.

liquid, CD that for the vapour. The portion BC , in which the pressure is constant and independent of the volume, corresponds to the process of vaporisation or condensation in which liquid and vapour are present simultaneously. BC represents the change of volume accompanying this change of state when the vapour is saturated, i.e., the liquid and vapour are in equilibrium with each

other. Increase of temperature increases the volume of the liquid but decreases the volume of the vapour, since in the latter case we have

$$v = R \frac{T}{p}$$

and the fraction T/p decreases with increasing pressure, as shown by the equation for the vapour pressure curve (49). The volume change becomes zero at M , which is the critical point at which the curves for v and v_1 meet without discontinuity.

We can imagine a continuous transition from liquid to vapour to take place along the partly unstable curve $ABCDEF$ in Fig. 4, Chapter III. If the minimum E lies below the abscissæ axis, the liquid can exist under a negative pressure.

(d) **The Vapour-Pressure Curve.** The vapour pressure increases with increasing temperature according to a curve having the

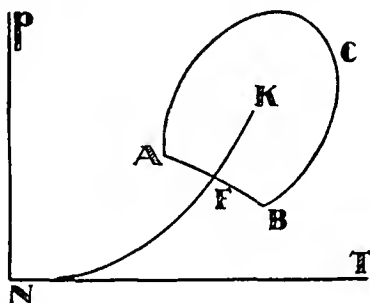


Fig. 3.

general form denoted by NK in Fig. 3. It runs from the absolute zero N to the critical point K , where the properties of liquid and vapour become identical. The transition from a point A (corresponding to the liquid state) to a point B (corresponding to the vapour state) can either take place by the direct path AFB , when evaporation (*i.e.*, a discontinuous change of state) will take place at F , or by the path ACB around the critical point, when the transition will be continuous along all parts of the curve. The latter possibility shows that there is no difference in principle between the liquid and vapour states. The phenomena of vaporisation and condensation can only take place at pressures and temperatures below the critical values.

The form of the vapour pressure curve can be derived by applying to the process of vaporisation the second law of thermodynamics in the form of equation I. (11),

$$q = T \left(\frac{\partial A}{\partial T} \right)_v,$$

q is here the heat of vaporisation and A the work of vaporisation.

If the two isothermals in Fig. 2 are only an infinitesimal distance apart, it is easily seen that dA in the above equation is equal to the area $BCC'B'$. We can thus write

$$dA = (v - v_1)dp, \quad (47)$$

where dp is the increase of vapour pressure which accompanies an increase of temperature dT . The dependence of the vapour pressure on the temperature is thus given by

$$q = T(v - v_1) \frac{dp}{dT}, \quad (48)$$

the so-called Clausius-Clapeyron equation. It is quite generally valid, since only the second law has been used in its derivation.

If we introduce the assumptions $v \gg v_1$ and $pv = RT$, (48) becomes

$$q = RT^2 \frac{d \ln p}{dT}, \quad (49)$$

from which it is seen that the relative vapour pressure increase is determined by the heat of vaporisation. The assumptions involved in (49) are however not rigorously true, and the equation is therefore only approximately valid.

If (49) is integrated on the assumption that q is constant over a certain temperature range, we obtain as an approximate expression for the vapour pressure curve in this temperature range

$$\ln p = -\frac{q}{R} \frac{1}{T} + I, \quad (50)$$

showing that there is a linear relation between $\ln p$ and the reciprocal of the temperature. q/R can be determined as the slope of the straight line which represents the relation between $\ln p$ and $1/T$.

If (6) is introduced into (49), we obtain

$$q - RT = RT^2 \frac{d \ln c}{dT},$$

or

$$\Delta E = RT^2 \frac{d \ln c}{dT}. \quad (51)$$

An equation analogous to (49) thus represents the relation between the relative change in the concentration of the saturated vapour

and the energy of vaporisation, sometimes called the "internal latent heat of vaporisation."

The complete course of the vapour pressure curve will be considered in a later section [X. 1. f.].

The temperature T at which the vapour pressure is p is called the "boiling point" of the liquid under a pressure p . If no pressure is specified, the boiling point is the value of T at normal atmospheric pressure.

(e) **The Equilibrium Crystal-liquid.** The melting point may be defined as the temperature at which the crystalline and liquid forms of a substance are in equilibrium. Above this temperature the crystals will melt completely, and below this temperature the

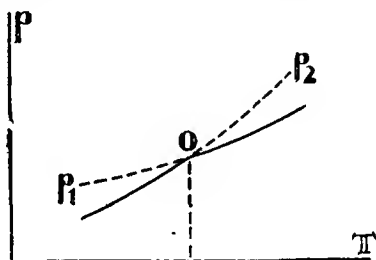


Fig. 4.

reverse process will take place spontaneously (the pressure being kept constant).

At the melting point the crystal and the liquid must have the same vapour pressure, since they are in equilibrium. At all other temperatures their vapour pressures will be different. The vapour pressure curve for the crystal is determined by an equation analogous to (48), q being now the heat of sublimation, i.e., the heat of vaporisation of the crystal. This must be greater than the heat of vaporisation of the liquid, since it may be regarded as the sum of the latter quantity and the heat of fusion. The vapour pressure curve of the crystal will therefore be steeper than that of the liquid. If the vapour pressures of the two phases are plotted in a p - T diagram, two curves will be obtained intersecting at the melting point (Fig. 4).

Since the form which has the greater vapour pressure must be unstable with respect to the other form, at temperatures below O in Fig. 4 the lower curve must refer to the crystal. This argument provides another proof that the curve for the crystal is the steeper. The part of the curve Op_2 cannot be realised in practice, while the part Op_1 is usually realisable.

The *work* done during fusion is normally very small compared with the work of vaporisation, on account of the small volume change which accompanies fusion. This work only becomes appreciable if fusion takes place at high pressures.

The *maximum work* which can be derived from the process of fusion (or crystallisation) is very nearly identical with the maximum work which can be obtained by transferring the vapour of the substance from the pressure p_1 to the pressure p_2 (or *vice versa*), where p_1 and p_2 are the vapour pressures of the crystal and the liquid respectively. We have therefore to a good approximation from (13)

$$A = \pm RT \ln \frac{p_1}{p_2}. \quad (52)$$

At the melting point $p_1 = p_2$, and the maximum work is therefore practically zero. Equation (52) is generally applicable to transfer from one condensed phase to another.

Since the work done is small, the *heat of fusion* will be very nearly equal to the energy supplied during fusion, and this equality will only cease to hold at high pressures. The relation between the heat of fusion and the vapour pressure curves of the liquid and the solid can be found in the following way. Just as the conversion of a gas from pressure $p + dp$ to pressure p can produce an amount of work vdp (equation (11)), so the conversion of a liquid of vapour pressure $p + dp$ to a solid of vapour pressure p can produce the same amount of work vdp . This work is zero at the melting point. At a temperature differing from the melting point by dT , the vapour pressure of the liquid is $p + dp$, and that of the crystal $p + dp_2$. The difference in vapour pressure is hence $dp_2 - dp_1$, and the conversion of crystal to liquid can provide an amount of work

$$dA = v(dp_2 - dp_1).$$

Introduction of equation I. (11) gives

$$q = Tv \left(\frac{dp_2}{dT} - \frac{dp_1}{dT} \right), \quad (53)$$

independent of whether the gas laws are valid for the vapour. This equation relates the heat of fusion of a given amount of substance to its volume in the vapour state and the direction of the vapour pressure curves of the solid and liquid phases at the melting point.

The application of equation (45) to the process of fusion gives directly

$$\frac{dq}{dT} = c_1 - c_2, \quad (54)$$

where c_1 and c_2 are the specific heats of the liquid and the crystal respectively. Since $c_1 > c_2$, the numerical value of the heat of fusion will increase with increasing temperature.

(f) **The Effect of Pressure on the Melting Point.** By applying the equation $q = T \frac{dA}{dT}$ to the process of fusion we obtain (just as for vaporisation)

$$q = T(v_1 - v_2) \frac{dp}{dT} \quad (55)$$

This equation expresses the relation between the heat of fusion and $\frac{dT}{dp}$, i.e., the variation of the melting point with pressure. q is positive, as in the case of vaporisation. On the other hand, while the increase of volume on vaporisation is always positive and relatively large, $v_1 - v_2$ for fusion is usually small and may be either positive or negative. If $v_1 - v_2$ is positive, i.e., if the density of the crystal is greater than that of the liquid (as is usually the case), then $\frac{dp}{dT}$ must also be positive and the melting point will be increased

by pressure. In the case of water and a few other substances $v_1 - v_2$ is negative, and the melting point will therefore decrease with increasing pressure.

It is of course possible to use equation (55) quantitatively for calculating q or $\frac{dT}{dp}$. If the volume is expressed in litres and the pressure in atmospheres, then the right-hand side of (55) will be expressed in litre-atmospheres and the heat of fusion must be expressed in the same units. Thus for 1 gram of water $q = 80.3$ cal. = $80.3/24.20$ l.atm., $v_1 = 0.00100$, $v_2 = 0.00109$, $T = 273^\circ$.

These values give $\frac{dT}{dp} = -0.0077^\circ$, while experiment gave a value of -0.0078° for an increase of one atmosphere in the pressure.

If equation (55) is integrated, taking into account the variation of q with temperature, a general expression for the "melting curve" is obtained. Experimental investigations with very high pressures and correspondingly high displacements of the melting point have been recently carried out (e.g., by Tammann).

(g) **Allotropic and Polymorphic Transformations.** From a thermodynamic standpoint, allotropic or polymorphic transformations do not differ from the processes of fusion and crystallisation. If the vapour pressure curves of two allotropic forms are plotted in a p - T diagram we obtain two curves intersecting at the tempera-

ture at which the two forms can exist simultaneously in equilibrium with each other, i.e., the *transition temperature*. In Fig. 5 this point is marked as *O* and the two curves II and III. Below *O* (i.e., at temperatures below the transition point) the form III has the smaller vapour pressure and is thus the stable form, while II will be the stable form above *O*. If I represents the vapour pressure curve of the molten substance, then (as shown above) *P* and *Q* are the melting points of the two forms, and it will be seen from the figure that the melting point of the unstable form III must be lower than the melting point of the stable form II. If the position of the vapour pressure curve of the liquid phase is as shown by I' in the figure, then it will never be possible to realise the transition point *O*, and III will always appear as the more stable of the two crystal-

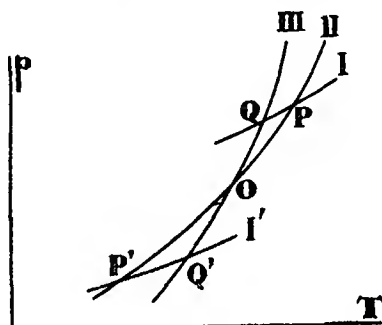


Fig 5.

line forms (monotropy or enantiotropy). There will be a similar state of affairs if I or I' represents the vapour pressure curve of a third solid modification of the substance.

These considerations presuppose that the vapours produced by the two forms are chemically identical. If this is so, we are dealing with polymorphism, or physical isomerism, while substances which are chemical isomers usually form vapours which are also isomeric. In the latter case the rules for vapour pressure and melting point given above cannot be applied.

The production of the form stable at the higher temperature from the other form must always be accompanied by absorption of heat: this follows directly from equation I. (10). The effect of pressure upon the transition temperature can be calculated from the heat of transition just as in the case of fusion.

It appears to be a generally valid rule that at a melting point or transition point the form which is stable at a higher temperature has

the larger specific heat. In the case of a melting point this means that the specific heat of the liquid form is greater than that of the solid. The numerical value of the heat of fusion or transition will therefore increase with increasing temperature.

(h) **Application of Thermodynamic Functions for Determining Phase Equilibrium.** Any of the thermodynamic equilibrium conditions given in the last chapter can be used as a criterion for the equilibrium between two phases of the same substances. Particularly simple relations are obtained by using the thermodynamic potential G .

In terms of this function the equilibrium condition is I. (35),

$$(dG)_{T,p} = 0,$$

where dG is the change in the thermodynamic potential of the two-phase system caused by the transfer of substance from phase (1) to phase (2). The equilibrium condition can thus be written

$$G_1 = G_2,$$

where G_1 and G_2 are the potentials of the two phases per unit quantity of substance.

If the temperature and pressure are altered, we have from I. (33)

$$dG_1 = -S_1 dT + v_1 dp,$$

$$dG_2 = -S_2 dT + v_2 dp,$$

where S and v also refer to unit quantity of substance. If equilibrium is to be maintained during this change of state, we must have

$$dG_1 = dG_2,$$

$$\text{and hence} \quad (S_1 - S_2) dT = (v_1 - v_2) dp, \quad (56)$$

$$\text{or} \quad T(S_1 - S_2) = T(v_1 - v_2) \frac{dp}{dT}, \quad (57)$$

where dp is the increase in equilibrium pressure which corresponds to the temperature increase dT . (56) exemplifies the simple relation which exists between the extensive and intensive factors of the forms of energy which play a part in displacement of equilibria.

Equation I. (14) gives

$$T(S_1 - S_2) = q, \quad (58)$$

showing that (57) includes the Clausius-Clapeyron equation for the vapour pressure curve and the analogous formula for the effect of

pressure on the melting point. (58) also shows the general relation between the entropy transferred and the heat absorbed for displacements of equilibrium.

The same results may be obtained by applying I. (56) to a one-component system, since in such systems μ is identical with the value of G per unit quantity of substance (cf. I. (49)).

(i) **The Phase Diagram.** If the equilibrium curves for all the

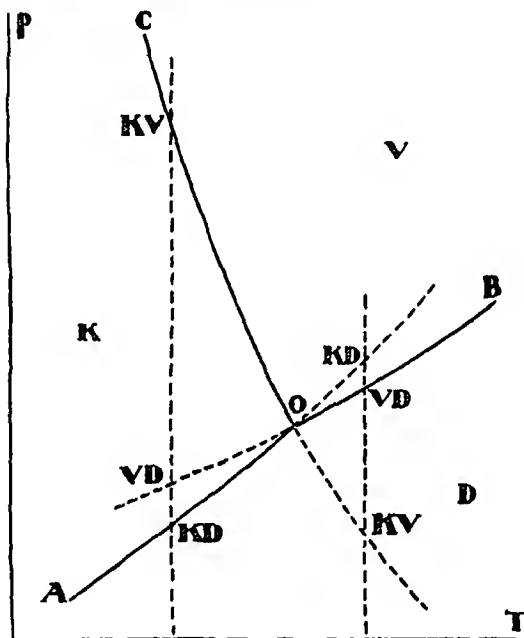


Fig. 6.

changes of state which a substance can undergo are plotted on a p - T diagram, we obtain the *phase diagram* of the substance. A typical diagram of this kind is shown in Fig. 6.

AO and OB are the vapour pressure curves of the solid and liquid forms of the substance respectively, and the point of intersection must therefore represent the melting point under the pressure of the saturated vapour. At the point O the solid, liquid and vapour are in equilibrium with each other, and the point must therefore also lie on the curve giving the variation of the melting point with pressure, represented by OC in the figure. Points on

the three equilibrium curves AO , BO and CO thus correspond to states in which the substance can exist in two phases, namely solid-vapour, liquid-vapour and solid-liquid respectively. These three curves all meet in the point O , at which all three states can co-exist, and which is hence known as the *triple point*. Since the co-ordinates of this point are completely fixed, its position can be regarded as a natural constant characteristic of the substance.

Points which do not lie on the curves represent a single state of aggregation. In the area between AO and OB only vapour can exist in a stable state: similarly between AO and CO we have solid and between BO and CO liquid. As already noted, the curve

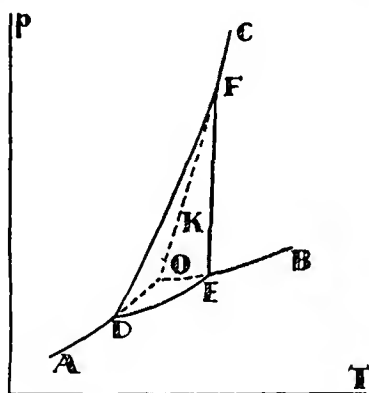


Fig. 7.

OC will in most cases slope to the right, since the melting point usually increases with increasing pressure. The diagram in Fig. 6 thus corresponds to a substance which is an exception to this rule, *e.g.*, water. The position of the triple point relative to the normal melting point at a pressure of one atmosphere depends on the pressure at the triple point and the direction of OC . It will be readily seen that for water the triple point is at a higher temperature than the melting point of ice under a pressure of one atmosphere.

If the solid substance can exist in various modifications the phase diagram will of course be more complicated, since there will be a triple point corresponding to each melting point or transition temperature. Fig. 7 represents a case in which the solid can exist in two forms, *e.g.*, sulphur.

AD , DE and EB are respectively the vapour pressure curves of rhombic, monoclinic and liquid sulphur. The points of intersection

D and *E* are respectively the transition point between the two allotropic modifications (95°) and the melting point of monoclinic sulphur (120°). These two points form the starting points of two new equilibrium curves, representing the effect of pressure upon the transition point and the melting point of monoclinic sulphur. The points *D* and *E* are thus triple points. A third triple point occurs at the intersection of the two last equilibrium curves at *F*, from which point another equilibrium curve *FC* proceeds. Finally, if the three curves *AD*, *EB* and *CF* are prolonged outside their stable region they will meet at *O*, the melting point of rhombic sulphur. This point is really unstable, but can be realised if rhombic sulphur is rapidly heated, when the curve *AO* can be extended past the point *D* without transformation to monoclinic sulphur taking place.

If the three equilibrium curves in Fig. 6 are prolonged through the triple point into the regions of instability, it will be seen that at an arbitrarily chosen temperature there are three possible equilibria, crystal-vapour (*KD*), liquid-vapour (*VD*) and crystal-liquid (*KV*), of which either one or two are unstable. If the diagram is used as a basis for predicting stability, it will be seen to confirm the rule previously stated, that the condensed phase having the greater vapour pressure will be unstable. It will also be seen that no more general law can be stated about the instability of phase equilibria at high pressures, but that the following rule is of general validity.

If we consider two equilibria involving a common phase (this comprises the pairs *KV-KD*, *VK-VD* and *KD-VD*), then the equilibrium with the lower pressure will be stable if the common phase has the greatest specific volume, while the equilibrium with the higher pressure will be stable if the common phase has the lowest specific volume. If the common phase has a specific volume intermediate between the specific volumes of the other two phases, then either both equilibria will be stable, or both will be unstable.

CHAPTER III

MOLECULAR THEORY

1. THE BASIS AND AIMS OF MOLECULAR KINETIC THEORY

IN addition to thermodynamics, the laws and methods of which have been developed and applied in the previous chapters, we have in the molecular kinetic conception of matter an instrument of the greatest importance for interpreting physical and chemical phenomena. While thermodynamics is essentially based on experimental evidence as to the non-existence of perpetual motion of the first and second kinds, molecular theory is characterised by certain hypothetical assumptions from which it is possible to deduce laws susceptible to experimental test. The aim of molecular kinetic theory is to account for the thermodynamic and energetic behaviour of matter in terms of the chemical conception of matter as consisting of molecules and atoms, using the mechanical laws derived from a study of the equilibrium and motion of larger bodies. Its scope is however wider than that of thermodynamics, since kinetic theory can also be used to interpret phenomena involving the time, *e.g.*, diffusion and rate of chemical reaction. It should be added that a detailed study of various phenomena has extended molecular theory beyond its original classical mechanical basis, and has led to new and important discoveries in the general laws governing the interaction of matter and energy.

(a) **The Avogadro Number. Atomic and Molecular Dimensions.** Our most general knowledge of the masses of atoms and molecules is expressed by the value assigned to the *Avogadro number* (also known as the Loschmidt number). This is defined as the number of actual molecules present in 1 gram-molecule; this number must be independent of the chemical nature of the substance considered, and is thus of universal application.

The Avogadro number, N_0 , has been determined by numerous methods, some of which will be dealt with later. Probably the most accurate value has been obtained by measuring the effect of an electric charge on the rate of fall of small drops of oil, suspended in air, in an electric field (Millikan). In these experiments it was found that there was a *minimum charge effect*, together with other effects which were multiples of the minimum effect. This behaviour

can only be accounted for by assuming that the drop carries an *elementary quantum of electricity*, or a discrete number of such quanta. The size of this elementary charge ϵ could be determined from its effect on the rate of fall in a field of known strength, giving the value $\epsilon = 4.774 \times 10^{-10}$ electrostatic units. Since further the amount of electricity F held by 1 gram-molecule of a substance is known [IX. 1 b], we can now calculate the number of elementary charges in the charge carried by 1 gram-equivalent. This gives as the most probable value

$$N_0 = \frac{F}{\epsilon} = 6.06 \times 10^{23}. \quad (1)$$

The Avogadro number can be used to obtain an approximate value for molecular dimensions, *e.g.*, by assuming spherical molecules which are cubically packed in the condensed state. If r is the molecular radius and v the volume of 1 gram-molecule, we have clearly

$$v = N_0 \cdot (2r)^3, \quad (2)$$

or

$$r \simeq 0.6 \sqrt[3]{v},$$

if v is expressed in c.c. and r in Ångström units ($1\text{Å} = 10^{-8}\text{ cm.}$). Thus if $v = 100$, we have

$$r \simeq 3\text{Å}.$$

Molecular dimensions calculated in this way are however somewhat too large, partly because it is incorrect to assume cubical packing in the condensed state.

Numerous other methods for determining r agree in giving values of about $1 - 3\text{Å}$ for the dimensions of atoms and simple molecules. By means of X-ray investigations the element of volume $(2r)^3$ can be determined very accurately in the case of crystals, leading to a value of $N_0 = 6.04 \times 10^{23}$ for the Avogadro number.

(b) The Equipartition of Energy. An important aspect of molecular theory is the assumption that the molecules are in continual thermal motion, characteristic of the substance and the state. According to the classical view, the usual mechanical laws valid for large particles also apply to the thermal motion of molecules, it being assumed in particular that the laws of conservation of energy and momentum can be used. Their application to molecular motion leads to the important law of *equipartition of energy*.

According to this law, if we have an arbitrary system of particles in motion, then when equilibrium has been reached the mean value of the kinetic energy of a particle for each *independent type of motion* will be a constant, independent of the nature, shape or mass of the particles.

The number of independent types of motion, or *degrees of freedom*, is thus the factor which determines the kinetic energy of a particle. This number may be defined as the number of components which are necessary and sufficient to describe the motion. For a dimensionless particle able to move in a straight line in space, the number of degrees of freedom is 3. If the particle is constrained to move in a plane, or along a fixed straight line, then the number of degrees of freedom will be 2 and 1 respectively, and the mean energy will be respectively $2/3$ and $1/3$ of the mean value for a particle free to move in space.

If the particle has finite dimensions in space, it will also be able to acquire rotational energy. Since rotational motion can be described in terms of three axes of rotation, the particle will acquire three additional degrees of freedom. In the case of a linear or one-dimensional particle, the number of rotational degrees of freedom will obviously be only two.

If the particles are not rigid, as assumed above, they will possess further degrees of freedom in virtue of the elastic vibration of their component parts. Such particles can acquire vibrational energy, in addition to the forms of energy described above, and will therefore have a higher energy content than particles of simpler structure.

It should be specially noted that the law of equipartition of energy deals with the *mean energy*, either for one molecule averaged over a sufficient period of time, or for a single instant of time averaged over a sufficiently large number of like molecules. The deviations of the velocity and energy content of molecules from the mean values are expressed by the *Maxwell distribution law*, dealt with below. We shall also see in the next section how the absolute value of the mean energy per degree of freedom can be calculated.

2. THE KINETIC THEORY OF THE FLUID STATE

(a) **Gas Pressure and the Kinetic Energy of Molecules.** According to the kinetic theory of gases (Clausius, Maxwell, Boltzmann) the pressure exerted by a gas on the walls of its containing vessel is attributed to the impact of the gas molecules. The magnitude of the pressure can be calculated by applying the ordinary laws of dynamics to the motion of the molecules.

If the mass of the molecule is m and its velocity w , we can write for its translational kinetic energy

$$E = \frac{1}{2}mw^2, \quad (3)$$

and for its momentum (or impulse)

$$i = mw. \quad (4)$$

If the molecule undergoes an elastic collision with the wall at right angles, it will undergo a change of momentum $2mw$, since after collision its velocity will be the same numerically, but opposite in sign. If the gas contains n molecules per c.c., and we imagine that $1/6$ of this number are moving at right angles to each of the six sides of a unit cube, then the momentum received by 1 sq. cm. per second will be

$$\frac{n}{6} w \cdot 2mw.$$

Since the pressure is equal to the rate of change of momentum per unit area, we have

$$p = \frac{1}{3}nmw^2. \quad (5)$$

Introducing for the weight of 1 c.c.,

$$nm = \frac{M}{v}$$

where M is the molecular weight and v the molecular volume in c.c., this becomes

$$pv = \frac{1}{3}Mw^2. \quad (6)$$

If we assume (as the basis of the theory demands) that the velocity of the molecules is independent of the volume, then the right-hand side of (6) is a constant, and the equation expresses *Boyle's law*.

According to the law of equipartition of energy, at thermal equilibrium the quantity E in equation (3) is independent of the nature of the molecules. We therefore have for different kinds of molecules 1, 2 . . . at constant temperature.

$$M_1w_1^2 = M_2w_2^2 = \dots, \quad (7)$$

showing that at constant temperature and pressure, v , the molar volume of the gas, is also independent of the nature of the molecules. This expresses *Avogadro's law*.

The effect of temperature on the pressure and molecular kinetic energy of a gas cannot be deduced directly from the kinetic picture, but must be determined by introducing the definition given by the

gas thermometer scale. In other words, we must introduce the equation of state II. (2)

$$pv = RT.$$

Combining this with (6) gives

$$\frac{1}{3}Mw^2 = RT, \quad (8)$$

showing that the square of the molecular velocity is proportional to the temperature. Introducing (3) in the form

$$E = \frac{1}{2}Mw^2, \quad (9)$$

we obtain the following simple expression for the translational energy of 1 gram-molecule of a gas

$$E = \frac{3}{2}RT. \quad (10)$$

The translational energy of a single molecule is analogously given by

$$E = \frac{3}{2}kT, \quad (11)$$

and the energy per degree of freedom (since the translational energy has three degrees of freedom) by

$$E = \frac{1}{2}kT, \quad (12)$$

where

$$k = \frac{R}{N_0}, \quad (13)$$

is the gas constant per molecule, also known as the *Boltzmann constant*. Since the gas constant R in absolute units has the value 8.32×10^7 ergs per degree, the value of the Boltzmann constant in the same units is

$$k = 1.373 \times 10^{-16} \quad (14)$$

This value can be used to calculate from the above formulæ numerical values for the velocity and energy of molecules as functions of the temperature. The energy per degree of freedom of a single molecule is thus given by

$$E_1 = 0.687 \times 10^{-16} T \text{ ergs} \quad (15)$$

At ordinary temperatures the velocities of molecules of normal size are found to lie between 100 and 1,000 metres per second.

It was assumed in the deduction of the above formulæ that at a given temperature all like molecules have the same velocity and translational energy. This condition is not actually fulfilled in practice, but a more exact treatment, taking into account the actual distribution of velocities and energies, leads to formulæ identical with those derived above. In this case, however, the term w^2 is not to be interpreted as the square of a constant velocity w , but as the arithmetic mean of the squares of all the velocities actually

present. In the same way, E denotes the arithmetic mean of the translational energies of all the molecules.

(b) **The Distribution of Velocities and Energies. Maxwell's Law.** In dealing with many of the properties of the fluid state, the law governing the distribution of different velocities in a system of molecules is of little importance. However, there are certain problems, dealing both with kinetics and with equilibrium states, in which the fact that the velocities of the single molecules vary within wide limits is of decisive importance. The actual distribution of velocities can be calculated by a statistical treatment, and the result is expressed in *Maxwell's distribution law*. We shall describe here some of the fundamental principles of the theory and some of the results and formulæ which can be derived from it.

Let the components of the velocity of a molecule along the x -, y - and z -axes of a rectangular co-ordinate system be w_x , w_y and w_z respectively, and let the resultant velocities parallel to the x -axis, parallel to the x - y -plane, and in the x - y - z space be respectively w_x , w_{xy} and w_{xyz} . We then have the three equations

$$\left. \begin{aligned} w_x^2 &= w_z^2 \\ w_{xy}^2 &= w_x^2 + w_y^2 \\ w_{xyz}^2 &= w_x^2 + w_y^2 + w_z^2 \end{aligned} \right\} \quad (16)$$

The probability that a molecule will have a velocity component parallel to the x -axis lying between w_x and $w_x + dw_x$ can be expressed as $f(w_x)dw_x$, since the probability must depend in some way on the velocity w_x which the molecule is required to possess, and must also be proportional to the size of the interval dw_x . The same argument applies to the other two axes. Since the probabilities P must be independent of the direction in space, we can write

$$P_x = f(w_x)dw_x, P_y = f(w_y)dw_y, P_z = f(w_z)dw_z, \quad (17)$$

where $f(w_x)$, $f(w_y)$ and $f(w_z)$ are functions of the same form. On account of the very large number of molecules present in a macroscopic amount of substance, P can be assumed to be proportional to the actual number of molecules which at any instant have velocities in the prescribed interval. Then if N_0 is the total number of molecules, we have

$$P_x = \frac{dN_x}{N_0}, P_y = \frac{dN_y}{N_0}, P_z = \frac{dN_z}{N_0}. \quad (18)$$

The distributions of velocity along the three axes must be mutually independent. Hence if P_x is the probability that a molecule has a velocity component in the interval w_x to $w_x + dw_x$, P_{xy} the probability of simultaneous velocity components in the intervals w_x

to $w_x + dw_x$ and w_y to $w_y + dw_y$, and finally P_{xyz} , the probability of simultaneous velocity components in the three intervals w_x to $w_x + dw_x$, w_y to $w_y + dw_y$ and w_z to $w_z + dw_z$, then we have

$$\left. \begin{aligned} P_x &= \frac{dN_x}{N_0} = f(w_x) dw_x, \\ P_{xy} &= \frac{dN_{xy}}{N_0} = f(w_x) f(w_y) dw_x dw_y, \\ P_{xyz} &= \frac{dN_{xyz}}{N_0} = f(w_x) f(w_y) f(w_z) dw_x dw_y dw_z. \end{aligned} \right\} \quad (19)$$

The function $f(w_x)$ (and the analogous functions $f(w_y)$ and $f(w_z)$) can be derived from the expression for dN_{xy} in equation (19) in the following manner. Taking a rectangular co-ordinate system in two dimensions, a molecule in the gas considered can be represented by

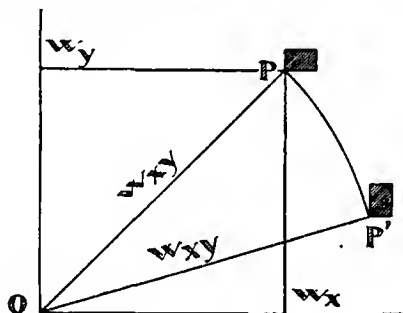


Fig. 1.

plotting its velocity components along the two axes; thus in Fig. 1 the point P represents a molecule having components of velocity w_x and w_y . The motion of this molecule in the xy plane is then represented by the line OP , the direction of which gives the direction of motion, and the length of which gives the velocity in this plane. By equation (16), the velocity w_{xy} is related to the component velocities by the expression

$$w_{xy}^2 = w_x^2 + w_y^2. \quad (20)$$

If every one of the N_0 molecules in the mass of gas considered is represented on the diagram in the same way, then the number of molecules dN_{xy} , having velocity components simultaneously between w_x and $w_x + dw_x$ and between w_y and $w_y + dw_y$ will be represented by points inside the shaded rectangle with sides dw_x and dw_y . This number of points is given by (19) as

$$dN_{xy} = N_0 f(w_x) f(w_y) dw_x dw_y, \quad (21)$$

and it can also be defined as the number of molecules having velocities along the x - and y -axes not less than w_x and w_y respectively, and whose velocities are also limited in the way indicated by the element of

area $dw_x dw_y$. Since the density of points is constant in the near neighbourhood of a given point, the number dN_{xy} will only depend upon the size of the element of area $dw_x dw_y$, and not on the separate linear intervals dw_x and dw_y .

The number of molecules dN_{xy} whose velocity components are not less than w_x and w_y and whose velocity is limited by the element of area $dw_x dw_y$ can also be defined as the number of molecules having a velocity not less than w_{xy} in the direction OP , and whose velocity is limited as before by the element of area. Since the velocity distribution is independent of the direction in the xy -plane, the density of points inside an element of area will depend only on the distance of the element from the origin O . If therefore we consider a second point P' the same distance from O as P , then the number of points in the element $dw_x dw_y$ will be the same when it is situated around P' as when it is situated around P . Hence as long as the condition of constant distance from the origin is fulfilled, i.e., as long as

$$w_x^2 + w_y^2 = \text{constant}, \quad (22)$$

a given element of area $dw_x dw_y$ will contain a constant number of molecules dN_{xy} , so that (21) gives

$$f(w_x)f(w_y) = \text{constant}. \quad (23)$$

The simultaneous validity of (22) and (23) leads directly to the function required. Differentiation of the two equations gives

$$w_x dw_x + w_y dw_y = 0 \quad (24)$$

and

$$d \ln f(w_x) + d \ln f(w_y) = 0. \quad (25)$$

Since the motions along the x -axis and along the y -axis are independent of each other, these two equations can only be simultaneously valid for different values of w_x and w_y provided that the ratio of the terms involving w_x is a constant independent of w_y , and the ratio of the terms involving w_y is a constant independent of w_x . We can thus write

$$d \ln f(w_x) = -\frac{2}{\alpha^2} w_x dw_x, \quad (26)$$

where $-\frac{2}{\alpha^2}$ is an arbitrary constant proportionality factor. Integration of (26) gives

$$\ln f(w_x) = -\frac{w_x^2}{\alpha^2} + \ln A, \quad (27)$$

where $\ln A$ is an arbitrary integration constant. Writing this in another form, the required function is given by

$$f(w_x) = A \cdot e^{-\frac{w_x^2}{\alpha^2}}. \quad (28)$$

In order to determine the value of the constant A , the expression for dN_x in (19) is integrated between the limits of velocity $+\infty$ and $-\infty$,

using (28). The left-hand side of the equation becomes unity, and hence

$$A \int_{-\infty}^{+\infty} e^{-\frac{w_x^2}{\alpha^2}} dw_x = 1.$$

On carrying out the integration we find

$$A = \frac{1}{\alpha\sqrt{\pi}},$$

so that (28) can be written

$$f(w_x) = \frac{1}{\alpha\sqrt{\pi}} e^{-\frac{w_x^2}{\alpha^2}}. \quad (29)$$

By introducing (29) and analogous equations for $f(w_y)$ and $f(w_z)$ into (19), we obtain the equations

$$\frac{dN_x}{N_0} = \frac{1}{\alpha\sqrt{\pi}} e^{-\frac{w_x^2}{\alpha^2}} dw_x, \quad (30)$$

$$\frac{dN_{xy}}{N_0} = \frac{1}{\alpha^2\pi} e^{-\frac{w_x^2}{\alpha^2}} dw_x dw_y, \quad (31)$$

$$\frac{dN_{xyz}}{N_0} = \frac{1}{\alpha^3\pi^{\frac{3}{2}}} e^{-\frac{w_x^2}{\alpha^2}} dw_x dw_y dw_z. \quad (32)$$

These equations give the numbers of molecules satisfying the conditions specified when formulating (19).

It is possible to introduce into equations (31) and (32) the velocities in two and three dimensions respectively, *irrespective of direction*. This is done by integrating (31) and (32) over all directions in a plane and in space respectively. The resulting expressions (including (30) re-written in an analogous form) are as follows :

$$\frac{dN_1}{N_0} = \frac{1}{\alpha\sqrt{\pi}} e^{-\frac{w_1^2}{\alpha^2}} dw_1, \quad (33)$$

$$\frac{dN_2}{N_0} = \frac{2}{\alpha^2} e^{-\frac{w_2^2}{\alpha^2}} w_2 dw_2, \quad (34)$$

$$\frac{dN_3}{N_0} = \frac{4}{\alpha^3\sqrt{\pi}} e^{-\frac{w_3^2}{\alpha^2}} w_3^2 dw_3, \quad (35)$$

where w_1 , w_2 and w_3 are the velocities in one, two and three dimensions respectively, as given by equation (16).

The condition that

$$\frac{dN_3}{N_0 dw_3}$$

should be a maximum is the same as the condition for a maximum for

$$e^{-\frac{w_3^2}{\alpha^2}} w_3^2,$$

and on investigating the expression in the usual way we find the condition to be $w_3 = \alpha$. The constant α introduced in deriving (26) is thus the velocity in space for which there is the greatest number of molecules within a given velocity interval, i.e., it is the most probable velocity for the motion of the molecules in space.

If (34) is integrated, using the general relation

$$\int e^x dx = e^x,$$

we find

$$\frac{N}{N_0} = e^{-\frac{w_2^2}{\alpha^2}}, \quad (36)$$

where N is the number of molecules out of the total number N_0 having a two-dimensional velocity greater than w_2 . The mean value of w_2^2 is given by

$$\overline{w_2^2} = \frac{1}{N_0} \int_0^\infty N dw_2^2,$$

and on introducing the value of N/N_0 given by (36), we obtain

$$\overline{w_2^2} = -\alpha^2 \int_0^\infty e^{-\frac{w_2^2}{\alpha^2}} = \alpha^2,$$

i.e., the square of the most probable velocity in space is equal to the mean of the squares of the velocities for motion in two dimensions. Further, if use is made of the fact that the mean energy is proportional to the number of degrees of freedom, we can write

$$\alpha^2 = \overline{w_1^2} = \frac{2}{3} \overline{w_3^2} = 2 \overline{w_1^2}. \quad (37)$$

Combination of this equation with equation (8) gives

$$\frac{1}{2} m \alpha^2 = kT. \quad (38)$$

Equations (33) to (35) derived above express Maxwell's distribution law. The meaning of this law is illustrated in the two diagrams Fig. 2 and Fig. 3, where the "distribution density" $dN/N_0 dw$ is

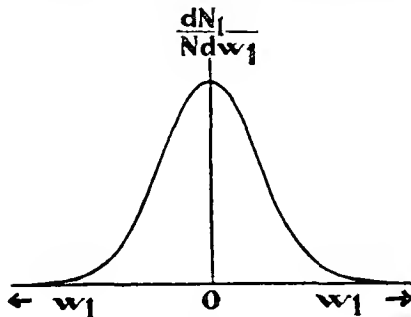


Fig. 2.

plotted as a function of w for motion in one and three dimensions respectively. Fig. 2 shows that the most probable velocity along a given axis is zero, while (as stated above) the most probable velocity in space is α . It will be seen that the number of molecules having velocities in excess of a given value is measured by that

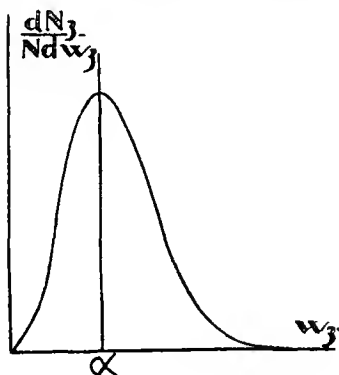


Fig. 3.

portion of the area between the curve and the w -axis which lies to the right of the given ordinate.

The distribution law can be expressed in terms of the translational energy of molecules in place of their velocities. This is done by

introducing (3) and (38) in equations (33) to (35). The simplest and most important of the expressions thus obtained is the one which expresses the distribution law in two dimensions, i.e.,

$$\frac{dN_2}{N_0} = \frac{1}{kT} e^{-\frac{F_2}{kT}} dE_2. \quad (39)$$

If we wish to find the number of molecules in a gas containing N_0 molecules which possess an energy (in two dimensions) of E_k or more, this number is easily obtained by integrating equation (39) between the limits E_k and ∞ . We then obtain (just as in the derivation of (36)) the following simple expression,

$$N = N_0 e^{-\frac{F_k}{kT}}. \quad (40)$$

This equation is of great importance in reaction kinetics.

(c) **Boltzmann's Distribution Law.** Consider a vessel divided into two parts, A and B , by two parallel planes close together, and let a constant field of force operate between the planes, tending to move all molecules between the planes in the direction $A \rightarrow B$.

On account of the presence of this field of force, each molecule in A can be said from the point of view of molecular theory to possess a *potential energy* [I. (2e)] ψ , the potential energy in B being taken as zero. Molecules passing through the field of force in the direction $A \rightarrow B$ will therefore receive an amount of kinetic energy ψ , while molecules passing in the opposite direction will lose the same amount of energy. For this reason there cannot be equilibrium when the molecules are evenly distributed through the vessel, but there must be a concentration of molecules in the space B . Let us assume that when equilibrium has been set up the numbers of molecules per c.c. in A and B are respectively N_A and N_B , where $N_B > N_A$.

According to equation (33), the number of molecules with velocities between w_A and $w_A + dw_A$ which leave A per square centimetre of the interface is given by

$$dN_A = N_A \frac{1}{\alpha \sqrt{\pi}} e^{-\frac{w_A^2}{\alpha^2}} w_A dw_A, \quad (41)$$

since this is the number which strikes the interface and passes into the space B . In order to maintain equilibrium in A , we must assume that an equal number of molecules in this velocity interval are supplied from B per second. These last molecules must however possess a higher velocity w_B in order to be able to overcome the field of force and then continue with velocity w_A . The number of

molecules in B which pass into A and replace the number (41) passing from A to B can be written in the form

$$dN_B = N_B \frac{1}{\alpha \sqrt{\pi}} e^{-\frac{w_B^2}{\alpha^2}} w_B dw_B \quad (42)$$

Since equilibrium is maintained, we thus have

$$N_A e^{-\frac{w_A^2}{\alpha^2}} dw_A^2 = N_B e^{-\frac{w_B^2}{\alpha^2}} dw_B^2. \quad (43)$$

Our assumptions about the field of force enable us to write

$$\frac{1}{2} m w_B^2 = \frac{1}{2} m w_A^2 + \psi, \quad (44)$$

where ψ is the constant field energy. ψ can also be written in the form

$$\psi = \frac{1}{2} m w_0^2, \quad (45)$$

where the new constant w_0 is the critical minimum velocity which the molecules in B must possess in order to be able to pass through the field of force. Equation (44) can then be written

$$w_A^2 = w_B^2 - w_0^2. \quad (46)$$

from which we obtain

$$dw_A^2 = dw_B^2. \quad (47)$$

Introducing (46) and (47) into (43), we have

$$\frac{N_B}{N_A} = e^{\frac{w_B^2 - w_A^2}{\alpha^2}} = e^{\frac{w_0^2}{\alpha^2}}, \quad (48)$$

or, using (38) and (45),

$$\frac{N_B}{N_A} = e^{\frac{\psi}{kT}}. \quad (49)$$

The last equation can also be written

$$\psi = kT \ln \frac{N_A}{N_B}, \quad (50)$$

or, if ψ is the potential energy for a whole gram-molecule,

$$\psi = RT \ln \frac{N_A}{N_B}. \quad (51)$$

These equations, which express Boltzmann's distribution law, are only strictly valid for a dilute gas, since no account has been

taken of molecular collisions on the volume occupied by the molecules themselves, both of which play a part in more concentrated systems.

In deriving the above formulæ, we have imagined an interface which brings about a difference of potential energy. It must be noted that the change in potential energy corresponding to the passage of molecules across this interface must always be identical with the change of potential energy associated with any other method of transferring the molecules from *A* to *B*, since if this were not the case it would be possible to produce a perpetual motion machine of the second kind.

The field of force acting can for example be a gravitational field. Equation (50) can then be used to calculate the density of a mass of gas at different heights above the earth's surface, ψ representing the difference in total potential energy corresponding to the difference in height. This leads directly to the well-known "hypsometric equation,"

$$M(\Phi_2 - \Phi_1) = Kl = kT \ln \frac{c_1}{c_2}, \quad (52)$$

where *K* is the force of gravity acting on a molecule, c_1 the concentration at the earth's surface, c_2 the concentration at height *l*, *M* the molecular weight, and Φ_1 and Φ_2 the two corresponding values for the gravitational potential. If we have, for example, a mass of gas in a closed cylinder with a vertical axis, it can be considered as being in a stationary state of diffusion in which the effect of kinetic tendency of the molecules to expand equally in all directions has been modified, partly by the force of gravity (which if acting alone would bring about complete sedimentation of all the molecules) and partly by the fixed limits which the walls of the vessel impose on the tendency to expand. This state of stationary equilibrium can be described formally as a state in which the effect of the thermal motion of the molecules is replaced by a force acting upon the molecules in the direction of diminishing concentration. The magnitude of this force is determined by the concentration gradient from equation (52), or, for a varying field, by the corresponding differential equation

$$K = -kT \frac{d \ln c}{dl}. \quad (53)$$

The importance of this conception lies in the use of the "force of diffusion" thus defined in treating states not in equilibrium, examples of which will be given later [IX. (2n).].

(d) **The Specific Heat of Gases. The Quantum of Energy.** If we

combine equation (10), expressing the translational energy of 1 gram-molecule of a gas, with the general equation I. (59) for specific heat,

$$\left(\frac{dE}{dT}\right)_v = c_v,$$

we obtain for the specific heat of a gas which obeys (10),

$$c_v = \frac{3}{2}R = 2.98. \quad (54)$$

We have further from II. (19)

$$c_p - c_v = R,$$

and hence

$$\frac{c_p}{c_v} = \frac{5}{3}, \quad (55)$$

and

$$c_p = 4.965. \quad (56)$$

Experimental determinations of c_p , c_v and the ratio c_p/c_v (which can be determined by some methods without determining the single specific heats) have shown that there is complete agreement with the above numerical values for certain gases such as helium, argon and mercury vapour. Equation (10), which was used in making the above calculations, only expresses the translational energy of the molecules, and it may therefore be concluded from the agreement between calculation and experiment that the energy of the gases mentioned above is exclusively translational. This agrees to a certain extent with the evidence from other sources that the molecules of these gases are monatomic, since it seems plausible to assume that a monatomic molecule cannot possess rotational or vibrational energy.

It might however be expected that the last two forms of energy would occur in diatomic molecules, and in support of this it is found that gases with diatomic molecules have higher specific heats. If such molecules are considered as one-dimensional, the rotational energy will have two degrees of freedom. Adding to these the three translational degrees of freedom and neglecting the possible occurrence of vibrational energy, this type of molecule possesses in all five degrees of freedom, so that

$$c_v = \frac{5}{2}R = 4.96, \quad (57)$$

$$\frac{c_p}{c_v} = \frac{7}{5} = 1.40, \quad (58)$$

$$c_p = \frac{7}{2}R = 6.95. \quad (59)$$

These equations are obeyed fairly well for a number of gases, *e.g.*, oxygen, nitrogen and carbon monoxide, while for other gases there are deviations, usually in a positive direction. Such deviations might be attributed to the fact that in deriving equations (57) to (59) no account was taken of the possible presence of vibrational energy: however, this explanation is not valid for differences of specific heat of less than $\frac{1}{2}R$, since the energy per degree of freedom is $\frac{1}{2}RT$. It is also important to note that although the specific heat is on the whole fairly constant over a considerable range of temperature, in many cases (particularly at low temperatures) a temperature dependence is observed which is not in agreement with the classical theory of specific heats which we have developed here. There is in fact no room for temperature dependence in this theory, since the equations contain no variable to which a variation of specific heat with temperature could be attributed. The classical theory has proved powerless to deal with this discrepancy, which was only explained by the development of new conceptions of the energetic behaviour of molecules, as expressed by the theory of energy quanta.

This theory was first proposed by M. Planck, and has played a decisive part in interpreting the relations between energy and matter. It supposes that in many cases the transference of energy is not continuous (as demanded by the classical theory), but quantised. This means that in the absorption of energy by a given species of molecule, the energy is taken up in the form of *energy quanta*, the size of the quantum being determined by the nature of the molecules under consideration. Thus if the quantum of energy is ϵ , the form of energy in question can only be taken up in quantities equal to ϵ or an integral multiple of ϵ .

The restrictions of the quantum theory apply primarily to the rotational and vibrational energy of molecules. That part of the energy of a gas which is due to the translational motion of the molecules is not affected by the theory, and it is for this reason that the specific heat of a monatomic gas (which possesses only translational energy) can be calculated by the classical methods. If, on the other hand, the energy can only be taken up in quanta, as for example in the case of rotational energy, the position is quite different. If the temperature is so low that the mean value of the molecular energy per degree of freedom, calculated from the

classical expression, is small compared with the quantum of energy, then rotational energy will not be acquired at all, and the molecules will not execute any rotational motion. As the temperature is increased and the classical mean value approaches the magnitude of the energy quantum, an increasing number of molecules will acquire a rotational quantum and the specific heat of the gas will thus rise. When the temperature has become so high that all the molecules are in possession of a large number of quanta, then the behaviour will again approach that of a classical system.

These considerations show that in applying the quantum theory to calculate specific heats, the decisive factor is the ratio between the temperature and the magnitude of the energy quantum. If the energy quantum is small, the transfer of energy will be nearly continuous, and deviations from the classical expressions will only appear at very low temperatures. If on the other hand the quanta are large, the departures from classical behaviour will be apparent at relatively high temperatures. As an example we may take hydrogen, which at ordinary temperatures has a specific heat of $c_p = 4.96$ (in agreement with the diatomic gases mentioned above), but which on lowering the temperature gives values approaching 2.98, the latter value being reached at 60°K . Further decrease of temperature causes no further change in the specific heat. This is a clear example of a diatomic molecule which at high temperatures has developed the five degrees of freedom corresponding to free translation and rotation, but which on cooling loses all its rotational energy, and from an energetic point of view becomes identical with a monatomic gas.

It should however be mentioned that a quantitative interpretation of the specific heat curve of hydrogen is only possible when allowance is made for the fact (predicted by the new quantum mechanics) that the hydrogen molecule can exist in two forms, ortho- and para-hydrogen, present in ordinary hydrogen in the ratio 3 : 1.

Similarly, chlorine is an example of a gas with a specific heat which exceeds the value $\frac{5}{2}R$ even at ordinary temperatures, and which increases with increasing temperature: this is due to the quantised development of new degrees of freedom corresponding to vibration of the atoms within the molecule.

The determination of the size of the energy quantum, and a more quantitative calculation of the distribution of quantised energy will be dealt with under the kinetic theory of the crystalline state.

(b) The Molecular Theory of the Liquid State. The forces of intermolecular attraction are only effective over very small distances which are not much greater than the dimensions of the molecules themselves.

It is therefore possible to treat gases in a dilute state without taking these forces into account, since the time during which the molecules are close together is small compared with the time during which they are far apart. In the liquid state, on the other hand, the distances between the molecules are always so small that the forces between the molecules become of decisive importance, in fact the stability of the liquid state is due to the existence of these forces.

The property which has been previously mentioned under the name of *cohesion*, and which was assumed to be responsible for the surface tension of liquids [II. 2], is a directly observable consequence of the intermolecular forces. The surface tension of a liquid is a force acting parallel to the tangent plane of the surface, in spite of the fact that the resultant of the cohesive forces acts at right angles to this plane; this is usually explained in the following way. If the extent of the surface is increased (which can be brought about by a change in the shape of the mass of liquid), the number of molecules in the surface will be increased, i.e., molecules must pass from the interior of the liquid to the surface. Since this takes place in opposition to the molecular forces (which, as we have seen, act in the opposite direction), work is necessary in order to increase the surface, and there must hence be a tension in the surface which opposes an increase of surface. The thermodynamic properties of this surface tension have already been dealt with.

The forces acting on a molecule in the interior of a liquid are not orientated, and their resultant will therefore be zero. However, on account of the attraction exerted on the molecules at the surface, the interior of the liquid will be subjected to a pressure, the *cohesion pressure*, which has a very considerable value for most liquids, and which in general will increase with the closeness of the molecules. The pressure in the interior of a liquid caused by the molecular forces will not however increase without limit as the liquid is compressed, since besides the attractive forces between the molecules there also exist repulsive forces which take effect when the molecules approach one another very closely, and increase inversely proportional to a very high power of the intermolecular distance. The combined effect of these two kinds of forces determines the so-called *internal pressure*.

This pressure can also be defined thermodynamically by the equation of state, I, (78), since it can be identified with the quantity $(\delta E/\delta v)_T$, which obviously has the dimensions of pressure. For perfect gases

TABLE I. *Internal Pressure of Ethyl Ether*

p (atm.)	Π (atm.)
0	2650
500	2850
1000	2790
1500	2670
2000	2530
2500	2320

this quantity is zero, corresponding to the fact that in this case the molecular forces are negligible.

Table I shows the internal pressure Π in ethyl ether as a function of the external pressure p . It will be seen that Π passes through a maximum in the neighbourhood of $p = 500$ atmospheres.

The effects produced by intermolecular forces in liquid will also be present to a certain extent in gases outside the range of ideal behaviour. If the pressure is not too great and the temperature not too high, the attractive forces will in general be the predominant factor. This means that the value of the product $p\bar{v}$ at constant temperature (which is a constant for perfect gases) will fall to a minimum with increasing pressure and then rise again.

Van der Waals has proposed an equation designed to describe the state of fluid substances taking into account both the forces acting between the molecules and the volume occupied by the molecules. In the first place, when the volume is decreased, the pressure will not increase so rapidly as predicted by Boyle's law, on account of the internal pressure produced by the attractive intermolecular forces. This internal pressure will increase rapidly with decreasing intermolecular distance, and according to van der Waals' theory can be put equal to

$$\frac{a}{v^2}.$$

In the second place, the theory takes into account the fact that the "free space" which is of importance in the equation of state is not equal to the measured volume of the gas, but is less than this on account of the volume of the molecules themselves. If the volume correction is put equal to b , we obtain the well-known *van der Waals' equation*,

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT. \quad (60)$$

By differentiating (60) at constant volume we obtain

$$\left(\frac{dp}{dT}\right)_v = \frac{R}{v - b}, \quad (61)$$

which in combination with (60) and I. (78) gives

$$\frac{a}{v^2} = \left(\frac{\partial E}{\partial v}\right)_T.$$

Thus, if equation (60) is obeyed, the quantity $\frac{a}{v^2}$ can be identified with the "internal pressure."

(f) **The Transition from Liquid to Vapour.** It has already been pointed out [II. (4c)] that the transition from liquid to vapour can

take place continuously. The van der Waals' equation, or any other equation of state which is based on similar theoretical treatment, leads to the same conclusion. The relation between pressure and volume at constant temperature expressed by equation (60) is shown graphically by the continuous curve in Fig. 4. However, this curve does not represent a stable state over its whole length, and in particular the portion *CDE* is characterised by an absolute in-

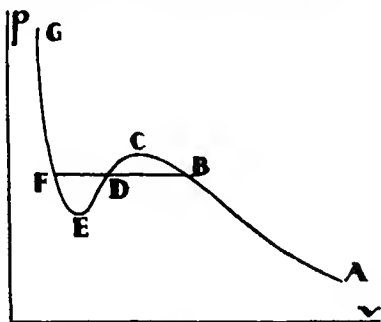


Fig. 4.

stability, since here an increase in the volume of the substance is accompanied by an increase of pressure, and *vice versa*.

As the state of the system passes from *A* (which may be taken to represent the perfect gas state) through *B* to *C*, the forces of attraction will increase with decreasing volume and become more and more important. When the point *C* is reached the state of the system is such that any increase of internal pressure produced by a further decrease of volume is just equal to the simultaneous increase of the thermal molecular pressure. Along *CE* the first of these pressures is greater than the second, but the state can only exist if all the molecules (which now have a great tendency to approach one another) maintain exactly the same distance from all neighbouring molecules, being held in this position of balance by exactly equal forces of attraction on all sides. Such a state is impossible to maintain on account of the thermal motion of the molecules, and in this region a transition takes place to the stable phase equilibrium liquid-vapour. The stable isotherm *ABFG* is characterised by the horizontal portion *BF*.

From a kinetic point of view, this phase equilibrium cannot be characterised as an immobile state in which nothing takes place, but as a state in which the number of molecules passing from the liquid to the vapour in unit time is equal to the number of molecules passing in the reverse direction. Since the surface of the liquid can be considered as a field of force in which the molecules experience a one-sided attraction towards the bulk of the liquid, it is natural

to assume that all molecules in the vapour phase which strike the surface will pass through it and become "condensed," independent of the velocity with which the collision takes place. On the other hand, if molecules in the liquid attempt to force their way out into the vapour, the retarding effect of the surface field of force will prevent a large proportion of them from "evaporating," since it will retain those molecules whose transitional kinetic energy in a direction at right angles to the surface does not exceed ψ , the potential energy characteristic of the field of force. This potential energy corresponds exactly to the potential energy difference which we assumed in paragraph (c) when dealing with uneven distribution between two gas phases.

The molecules which succeed in passing through the surface in opposition to the field of force will of course lose an amount of energy equal to ψ , while molecules passing through the surface in the reverse direction will gain an equal amount of energy. Vaporisation will therefore bring about a fall in temperature, and conversely the condensation of molecules from the vapour causes a rise in temperature. These temperature changes are occasioned by the difference between free and latent heat previously dealt with [I. (2b)]. The amount of energy which must be supplied or removed in order to keep the temperature constant while these processes take place is the "internal heat of vaporisation," *i.e.*, the total heat or work of vaporisation, minus the work of vaporisation. The above considerations show that the value of the internal heat of vaporisation is equal to the difference in molecular potential energy [I. (2e)] between the two phases.

We can also obtain expressions for the velocity of vaporisation of a liquid at constant temperature and for the equilibrium between liquid and vapour. The number of molecules with velocity components at right angles to the surface between w and $w + dw$ which strike 1 sq. cm. of the surface per second is given by equation (47) as

$$dN = N_B \frac{1}{\alpha \sqrt{\pi}} e^{-\frac{w^2}{\alpha^2}} w dw,$$

where N_B is the number of molecules per c.c. If w_k is the minimum velocity which allows the molecules to penetrate the surface, then integration of the above expression between the limits w_k and ∞ gives the total number of molecules penetrating the surface per second. This integration gives

$$N = N_B \frac{\alpha}{2\sqrt{\pi}} e^{-\frac{w_k^2}{\alpha^2}}, \quad (62)$$

or, from (38),

$$N = N_B \frac{\alpha}{2\sqrt{\pi}} e^{-\frac{E_k}{kT}}, \quad (63)$$

where E_k is the critical translational energy in one dimension corresponding to the velocity w_k . This gives the rate of evaporation, expressed as the number of molecules leaving 1 sq. cm. of liquid surface in 1 second.

The equilibrium expression is obtained by equating this number to the number of vapour molecules which strike the surface in the same time. If N_A is the number of molecules per c.c. in the vapour, the required number is obtained by integrating the expression

$$dN = N_A \frac{1}{\alpha\sqrt{\pi}} e^{-\frac{w^2}{\alpha^2}} w dw.$$

between the limits 0 and ∞ , since we have seen that all vapour molecules striking the surface are condensed. The integration gives

$$N = N_A \frac{\alpha}{2\sqrt{\pi}}, \quad (64)$$

and combination of (63) and (64) gives the equilibrium condition

$$\frac{N_B}{N_A} = e^{\frac{E_k}{kT}} \quad (65)$$

The form of this equation is identical with that of equation (49), which expresses the Boltzmann distribution law, and it could have been written down by direct analogy with this law, since the critical energy E_k which liquid molecules must possess in order to penetrate the surface corresponds exactly to the potential energy ψ appearing in equation (49).

Like (49), (65) cannot be considered an exact equation, since among other things the volume occupied by the molecules themselves was not taken into account in its derivation. An expression which agrees much better with experiment is obtained by first putting

$$\frac{N_B}{N_A} = \frac{v_A}{v_B},$$

where v_A and v_B are the molar volumes of the vapour and liquid respectively, and then correcting these volumes for the volume occupied by the molecules themselves, using the quantity b in van der Waal's equation. This gives

$$\frac{v_A - b}{v_B - b} = e^{\frac{E_k}{kT}}. \quad (66)$$

III. 3. THE KINETIC THEORY OF CRYSTALS 83

If (65) or (66) is differentiated with respect to temperature, assuming that E_k and N_B or $v_B - b$ are constant, and that $v_A \gg b$, we obtain

$$E_k = kT^2 \frac{d \ln N_A}{dT}, \quad (67)$$

which is identical with the thermodynamic vaporisation equation II. (51), if we put

$$N_0 E_k = \Delta E.$$

This means that in agreement with the above considerations we must identify the heat of vaporisation with the increase of potential energy which the molecules undergo when they pass from the liquid to the vapour through the surface field of force.

Equation (65) predicts that the ratio between the concentrations of molecules in the vapour and in the liquid should depend only on E_k/T . In agreement with this it is found that the ratio between the molar heat of vaporisation and the temperature at which the vapour possesses a given concentration is a fairly constant quantity. A related empirical rule known as *Trouton's rule* may be written

$$\frac{q}{T} = \text{constant} = \text{about } 21,$$

where q is the molar heat of vaporisation and T the boiling point under a pressure of one atmosphere. This relation can also be considered as expressing the constancy of the entropy of vaporisation under the conditions stated. It is, however, only approximately valid.

3. KINETIC THEORY OF THE CRYSTALLINE STATE

(a) **Lattice Structure and Lattice Energy.** The external geometrical form possessed by crystals is a consequence of an internal molecular structure differing essentially from the structure characterising the fluid state. While in the latter state the molecules can move freely, the constituent particles in a crystal form a so-called *crystal lattice*, in which each individual component has its fixed position, and the arrangement of the components follows simple geometrical laws. The thermal molecular motion of each single particle is confined to oscillations about a fixed position of rest.

The properties of a gas can in general be characterised as the properties of molecules, but in the case of crystals the molecule does not as a rule play such an important part. The particles which are arranged in the lattice can often be regarded as atoms or ions. The properties of crystals are affected to a high degree by the nature of the components of which the lattice is built up. In

investigating the structure of the crystal lattice, which is a very important factor in determining the properties of the crystal, the power of the crystal to diffract X-rays has proved particularly valuable. In certain respects the behaviour of ionic lattices is specially simple. Typical salts, *e.g.*, NaCl, are completely dissociated into ions in the crystalline state, and their properties can be calculated to some extent from the structure of the crystal and the laws for electrostatic forces. Thus it is possible to calculate the *lattice energy* of the crystal, *i.e.*, the amount of energy which must be supplied in order to turn the crystal into a completely ionised dilute vapour. The same quantity can be derived from the heat of formation of the crystal from the elements concerned and the ionisation energy of the latter [X. (2a)]. In the simplest cases there is good agreement between the two sets of values.

(b) **Specific Heat. Calculation from the Quantum Theory.** Since the thermal energy of a crystal depends on the vibrations of the particles in the crystal lattice about their equilibrium positions, it is reasonable to begin our calculation of the energy and the specific heat by considering a linear oscillator, *i.e.*, a particle constrained to oscillate in one dimension.

According to the law of equipartition of energy (12), such a linear oscillator must have the average kinetic energy $\frac{1}{2}kT$. It will, however, also possess a certain potential energy, which increases with its displacement from the point of equilibrium. If the oscillation is *harmonic*, *i.e.*, if the force acting on the particle is proportional to this displacement, the mean values of the potential and kinetic energies will be the same. A linear harmonic oscillator therefore has the energy

$$E = kT, \quad (68)$$

i.e., a total energy corresponding to two degrees of freedom.

Since in the crystal the particle can oscillate in three dimensions, and will possess the above amount of energy for each dimension, the total energy will be

$$E = 3kT \quad (69)$$

or

$$E = 3RT \quad (70)$$

for each gram-atom. This gives for the specific heat at constant volume

$$c_v = 3R = 5.96.$$

Since according to equation I. (68) c_p is somewhat greater than c_v , the above result agrees well with the rule of Dulong and Petit,

according to which the specific heat at constant pressure is somewhat greater than 6. However, the theory provides no explanation of the existing exceptions to this rule, or of the decrease of specific heat with decrease of temperature which solids exhibit in general [II. 3]. Such an explanation is only possible in terms of the quantum theory, which, as we have seen, is able in the case of gases to account at least qualitatively for the deviations from the classical laws.

A quantitative calculation of the energy content and specific heat of simple crystals was first carried out by Einstein for a crystal in which the units constituting the lattice are atoms. We shall give here a simplified derivation of Einstein's formula.

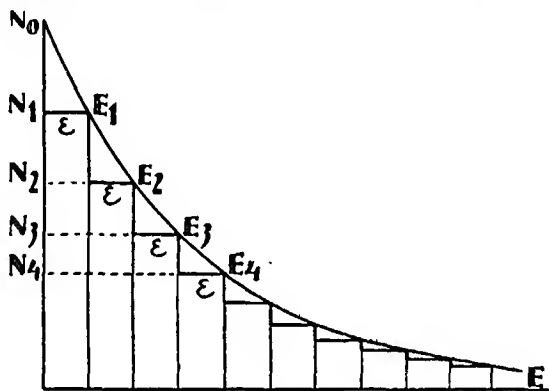


Fig. 6.

According to Maxwell's distribution law, the distribution of energy in a system with two degrees of freedom is given by equation (39)

$$\frac{dN}{N_0} = \frac{1}{kT} e^{-\frac{E}{kT}} dE.$$

If N_0 is the Avogadro number, dN is the fraction of the atoms in a gram-atom of crystal which have energy in two dimensions between E and $E + dE$. By integrating between E and ∞ we obtain in agreement with (40),

$$N = N_0 e^{-\frac{E}{kT}}. \quad (71)$$

Equation (71) is represented graphically by the curve in Fig. 5, in which the abscissa represents the energy E of a molecule, and the ordinate the number of molecules N having energy greater than E . The curve can also be obtained by imagining all the N_0 molecules

arranged at equal intervals along the ordinate axis in order of increasing energy, and then marking out an abscissa equal to the energy corresponding to each point. Since the equation can be written in the form

$$\ln \frac{N}{N_0} = - \frac{E}{kT} \quad (72)$$

this "distribution curve" must be a logarithmic curve of the type met with in other connections, *e.g.*, in the kinetics of a unimolecular process [VIII. (1c)]. It is easily seen from equation (72) that that part of the curve lying to the right of an arbitrary abscissa value will be a distribution curve for the number of molecules given by the corresponding ordinate. Boltzmann's distribution law can be derived directly from these considerations.

The total energy of the molecules is obviously represented by the area under the distribution curve and can be written as

$$E = \int_0^{\infty} N dE, \quad (73)$$

or, introducing (71),

$$E = - N_0 \int_0^{\infty} e^{-\frac{E}{kT}} dE = N_0 kT = RT. \quad (74)$$

This expression agrees with the classical value for the energy in two dimensions of 1 gram-molecule, as of course it must on account of the method by which it was derived.

According to the quantum theory, however, (74) does not give the correct value for the energy, since it assumes a continuous energy distribution from zero to infinity. The quantum assumption is that the oscillating atoms can only possess integral multiples of the quantum of energy ϵ , and further, that all atoms which according to Maxwell's law should have energies between 0 and ϵ are actually without energy, while atoms which according to Maxwell's law should have energies between ϵ and 2ϵ actually have only one quantum, etc. On this assumption the total energy will be represented by the area under the stepped line in Fig. 5, where the length of the horizontal steps is ϵ . This area is seen to be

$$E_t = \epsilon(N_1 + N_2 + N_3 + \dots), \quad (75)$$

where according to (72),

$$N_1 = N_0 x, \quad N_2 = N_1 x, \quad N_3 = N_2 x \dots, \quad (76)$$

and

$$x = e^{-\frac{\epsilon}{kT}}.$$

Combining these equations gives

$$E_t = N_0 \epsilon (x + x^2 + x^3 + \dots)$$

or, since the sum of the series inside the brackets is $\frac{x}{1-x}$,

$$E_t = N_0 \frac{\epsilon}{e^{kT} - 1}. \quad (77)$$

The factor multiplying N_0 in this formula is the energy (according to the quantum theory) of a particle with two degrees of freedom, i.e., the energy of a linear harmonic oscillator. In agreement with this, if the quantum of energy ϵ is sufficiently small, (77) becomes identical with (74), the denominator reducing to $\frac{\epsilon}{kT}$.

Since an atom oscillating freely in a crystal has three times as many degrees of freedom as a linear oscillator, the energy of one gram-atom of crystals will be

$$E = 3N_0 \frac{\epsilon}{e^{kT} - 1}, \quad (78)$$

an expression which reduces to the classical limiting value $3RT$, when $kT \gg \epsilon$, but which gives $E = 0$ at very low temperatures.

If (78) is differentiated with respect to temperature, we obtain for the specific heat

$$c_v = 3Re^{kT} \left(\frac{\frac{\epsilon}{kT}}{e^{kT} - 1} \right)^2, \quad (79)$$

from which the value of c_p can be calculated by the thermodynamic equations I. (68) and I. (69).

As regards the success of (79), Einstein's specific heat equation, it may be said to reproduce qualitatively the general features of the temperature variation of the specific heat of simple crystals. At a sufficiently high temperature it agrees with Dulong and Petit's law, while at a sufficiently low temperature it predicts a zero specific heat. It is not, however, in quantitative agreement with experiment, and in particular it does not agree with the simple relation found to hold at low temperatures, according to which the specific heat near absolute zero is proportional to the third power of the temperature [II. (43)].

Different substances behave very differently as regards the variation of their specific heats with temperature, thus showing

that the quantum of energy ϵ is not the same in all cases. A very important step in Planck's hypothesis of energy quanta lies in the relation postulated between the magnitude of the energy quantum and the frequency ν of the oscillator in question. This frequency is the number of oscillations which the oscillator executes per second, and for a simple harmonic oscillator is independent of the amplitude of oscillation. The relation postulated by Planck is

$$\epsilon = h\nu, \quad (80)$$

where h is a universal constant known as *Planck's action constant*, which has the dimensions energy \times time and the value in absolute units

$$h = 6.55 \times 10^{-27}. \quad (81)$$

Methods which serve to determine the characteristic frequency of an oscillator thus lead directly to a determination of the quantum of energy. We shall not, however, deal with these methods here.

One of the essential reasons for the partial failure of Einstein's specific heat formula lies in the fact that its derivation assumes that the oscillation of the atoms in a crystal is "monochromatic," i.e., is characterised by a single frequency. Actually, on account of the mutual interactions between the oscillations of the closely packed atoms we have to deal with a whole spectrum of oscillation frequencies. This has been taken into account by Debye, and leads to an equation differing somewhat from (79) which is in excellent agreement with experiment.

CHAPTER IV

MIXTURES

1. PURE SUBSTANCES AND MIXTURES

A HOMOGENEOUS mixture is a system whose properties are changed when a portion of the matter constituting the system passes into a new phase under arbitrarily chosen external conditions. A pure substance will remain unchanged when subjected to the same treatment under varying external conditions.

The formation of a new phase is most simply brought about by an ordinary change in the state of aggregation. Thus, for example, a pure substance like water undergoes no change of properties on partial freezing or partial vaporisation. Salt water, on the other hand, will leave behind a fraction with properties differing from those of the original sample, and is hence a mixture.

Certain mixtures will not give a fraction with different properties when subjected to a change of state under all external conditions. Thus a 20% solution of hydrochloric acid boils at 110° under a pressure of one atmosphere without undergoing any change in its composition. However, at any other temperature the mixed nature of solution will be apparent at this pressure.

These experiments do not, however, prove anything conclusively about the internal structure of the substances concerned, since it quite often happens that one kind of change of state indicates a substance to be pure, while another indicates it to be a mixture, *e.g.*, the freezing and vaporisation of "pure" sulphuric acid. In a case of this kind we speak of the dissociation of pure substances. The greater the range of conditions under which a substance appears to be pure, the greater is its stability.

From the point of view of molecular theory, any substance whose molecules are all alike can be termed a pure substance. It is not, however, always true that all the molecules in a pure substance are alike, *e.g.*, water probably contains both single molecules, H_2O and double molecules, H_4O_2 . Composite substances of this kind behave like pure substances during changes of state because the equilibrium between the different molecular species of which the substance consists is adjusted almost instantaneously as the change takes place.

Experience has shown that a mixture can always be split up into

(or formed from) at least two pure substances, which are known as the constituents or *components* of the system. The mixture is termed binary, ternary, quaternary, etc., according to the number of its components. Provided that no slow "chemical processes" take place on mixing, the properties of the mixture will be completely determined by the nature and relative amounts of the components, together with physical conditions such as temperature and pressure. For the present we shall deal only with mixtures of this kind, and shall confine ourselves chiefly to mixtures formed from only two components.

The composition of a mixture or solution can be defined by specifying the *concentration* in different ways. If the mixture contains n_1 molecules of the first component K_1 and n_2 molecules of the second component K_2 , the concentration can be specified by the ratio

$$\frac{n_1}{n_1 + n_2} = x. \quad \text{The concentration of } K_1 \text{ is then } x \text{ and that of } K_2$$

$1 - x$, so that the concentration scale extends from 0 to 1. x and $1 - x$ are termed the molecular or mole fractions of the components.

The concentration is also often expressed in terms of $\frac{n_1}{n_2}$, which

becomes equal to the mole fraction x in solutions which are extremely dilute with respect to K_1 . In the case of liquid mixtures the concentration can also be stated as the number of gram-molecules of K_1 in unit volume of the mixture, or as the number of grams per

unit volume or weight. In the first case we have $c = \frac{n_1}{V}$, or, if V'

is the molar volume (*i.e.*, the volume which contains one gram-molecule of the component in question), then $c = \frac{1}{V'}$. The molar

volume and the concentration thus have reciprocal values in this scale.

2. GASEOUS MIXTURES

(a) **Partial Pressures.** If n_1 molecules of a perfect gas are brought into contact with n_2 molecules of another perfect gas, the concentration of the mixture can be expressed by the mole fraction $\frac{n_1}{n_1 + n_2} = x$, and its properties can be calculated from the concentration by means of simple relations.

The pressure in a gaseous mixture can be considered as the sum of the *partial pressures* of the single components. These partial pressures are best defined by describing the method by which they can in principle be measured. This measurement is carried out by

placing the mixture in an enclosure R (Fig. 1) provided with two semi-permeable walls A and B , of which A is permeable to K_1 but

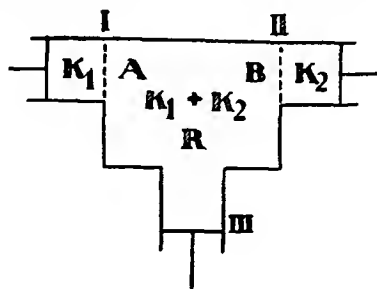


FIG. 1.

not to K_2 , while B is permeable to K_2 but not to K_1 . If the enclosure communicates through A and B with two compartments filled with frictionless pistons I and II, then the space between A and I will be filled with K_1 and the space between B and II with K_2 , while the pressures of the two gases can be measured directly by means of the pistons. These pressures, p_1 and p_2 , are the partial pressures of K_1 and K_2 respectively. They will be found to satisfy the following equations :—

$$\left. \begin{aligned} p_1 &= xp \\ p_2 &= (1-x)p, \end{aligned} \right\} \quad (1)$$

where p is the total pressure inside R . The partial pressure of each component is thus equal to the pressure which that component would exert if it alone occupied the space in question. This law can be easily deduced from the kinetic theory of gases.

(b) **Work of Mixing. The Thermodynamic Functions.** Let us assume that a vessel contains two gases K_1 and K_2 at the same temperature and pressure, separated by a partition. If the partition is removed spontaneous mixing will take place without any change in the total pressure. If the amounts of the gases in gram-molecules are x and $1-x$ and the original pressure p , then if the temperature is kept constant the partial pressures after mixing will be given by equation (1).

The maximum work which can be obtained from this spontaneous isothermal mixing process can be calculated in the following way. First let the pressures of the two components K_1 and K_2 be altered from p to p_1 and p_2 respectively. According to equation II. (10), the work thus obtained is given by

$$A = xRT \ln \frac{P}{P_1} + (1-x)RT \ln \frac{P}{P_2},$$

or, introducing (1),

$$A = -xRT \ln x - (1-x)RT \ln (1-x). \quad (2)$$

The two gases are then passed reversibly through the semi-permeable walls into the space R in Fig. 1, which already contains a mixture of concentration x . The work thus obtained is given by

$$A' = -xP_1v_1 - (1-x)P_2v_2 = -RT,$$

provided that the partial pressures in the mixture do not change while the gases are being introduced. The last condition can be fulfilled by introducing K_1 and K_2 continuously in the ratio $x : 1-x$, and at the same time removing the same quantities in the form of mixture by means of the piston III. Since the amount of mixture is one gram-molecule, the work done on the piston III is

$$A'' = RT.$$

The total work of mixing for the production of one gram-molecule of mixture is thus given by equation (2), since $A' + A'' = 0$.

Since equation (2) shows A to be proportional to the temperature, the energy change in the mixing process is given by I. (12) as zero. The heat evolution will hence also be zero when the mixing takes place in the normal way by diffusion of the gases without the production of any work.

If the process of mixing takes place isothermally and reversibly so that the work done is given by (2), then the same expression will represent the reversible heat absorption, q . The latter quantity is for reversible processes related to the entropy change by the equation

$$q = T\Delta S = T(S - S_0), \quad (3)$$

where S_0 is the entropy of the system before mixing and S the entropy after mixing. Hence, introducing (2) into (3),

$$\Delta S = S - S_0 = -R[x \ln x + (1-x) \ln (1-x)]. \quad (4)$$

The change which the other thermodynamic functions undergo on mixing is obtained by using I. (30), and remembering that ΔE is zero on account of (2) and I. (12). This gives

$$\Delta E = \Delta H = 0, \quad (5)$$

$$\Delta F = \Delta G = RT[x \ln x + (1-x) \ln (1-x)], \quad (6)$$

for one gram-molecule of mixture. (6) can, of course, also be derived directly from (2) and I. (32).

We can use equation (6) to calculate the dependence of the *chemical potentials* of the components upon the composition. It is simplest to use one of the definitions previously given in I. (45), i.e.,

$$\mu_1 = \left(\frac{\partial F}{\partial n_1} \right)_{T, v, n_2}, \quad (7)$$

$$\mu_1 = \left(\frac{\partial G}{\partial n_1} \right)_{T, p, n_2}. \quad (8)$$

It is therefore necessary to find the variation of F or G with n_1 , n_2 being kept constant. Since (6) refers to one gram-molecule of mixture, we shall use the expression

$$F - F_0 = G - G_0 = RT(n_1 + n_2)[x \ln x + (1 - x) \ln (1 - x)], \quad (9)$$

where the subscript 0 refers to the system *before* mixing. μ can then be found by differentiating (9) under the conditions expressed either in (7) or in (8). We shall use the second set of conditions for determining the potential.

In differentiating (9) with respect to n_1 , it should of course be noticed that both the terms inside the brackets are functions of n_1 . Carrying out a similar differentiation for μ_2 , we have

$$\left. \begin{aligned} \mu_1 - \mu_{1(1)} &= \left(\frac{\partial G - \partial G_0}{\partial n_1} \right)_{T, p, n_2} = RT \ln x, \\ \mu_2 - \mu_{2(2)} &= \left(\frac{\partial G - \partial G_0}{\partial n_2} \right)_{T, p, n_1} = RT \ln (1 - x), \end{aligned} \right\} \quad (10)$$

showing the simple relation between the chemical potentials and the concentrations. $\mu_{1(1)}$ and $\mu_{2(2)}$ are the potentials of the pure components at the temperature and pressure of the mixture.

It has been mentioned before that the absolute values of the thermodynamic functions are in general undetermined, since a standard state can be arbitrarily chosen. This will of course also apply to the potentials in equation (10). If we wish to emphasise this fact by writing (10) in the form

$$\left. \begin{aligned} \mu_1 &= RT \ln x + i_1, \\ \mu_2 &= RT \ln (1 - x) + i_2, \end{aligned} \right\} \quad (11)$$

then, in order to agree with (10), the integration constants i_1 and i_2 must be quantities independent of the concentration but depending on the temperature and pressure. In other words, if i_1 and i_2 are taken as constants, then equation (11) is valid only at constant temperature and pressure.

In general, the expressions for the thermodynamic functions of

gas mixtures can be readily derived from the expressions valid for pure gases by introducing the partial pressures in place of the total pressure. The correctness of this procedure follows from the independence of the energy on variations of pressure and concentration, coupled with the easily proved rule that the entropy of the gaseous system $n_1K_1 + n_2K_2$ has the same value when the two components are present in a given volume as a homogeneous mixture, and when they occupy this volume separately.

Thus the potential of K_1 in a gas mixture of varying pressure and composition is obtained from the equation valid for a pure gas, II. (32),

$$\mu_{1(1)} = RT \ln p + i,$$

by introducing the partial pressure p_1 , giving

$$\mu_1 = RT \ln p_1 + i. \quad (12)$$

By combining these two equations we obtain

$$\mu_1 - \mu_{1(1)} = RT \ln \frac{p_1}{p} = RT \ln x, \quad (13)$$

identical with equation (10).

3. LIQUID MIXTURES

While gases are miscible in all properties, liquids may be incompletely miscible if the intermolecular forces between the components K_1 and K_2 are much smaller than the forces between the molecules of the single components. In such cases, shaking the liquids together will form a heterogeneous liquid mixture or an emulsion. If the two components are completely miscible, it will be possible to make solutions of all compositions, from pure component K_2 , corresponding to $x = 0$, to pure component K_1 , corresponding to $x = 1$. It is found that the properties of solutions near $x = 0$ and $x = 1$ are characterised by specially simple relations: such "dilute solutions" are of particular importance and will be treated in the next chapter. In the present section we shall give the general theory of homogeneous mixtures without restrictions as to the concentration.

(a) **Work of Mixing.** Many of the properties of a homogeneous mixture can be derived if we know the relation between the concentration of the mixture and the chemical potentials of the components. It is not in general possible (as for mixtures of perfect gases) to express these potentials as known functions of the concentration, since their values also depend on the nature of the components.

There are, however, some general laws governing the effect of the concentration, which can be derived by considering the maximum work obtainable from the process of mixing.

This process can be carried out in two essentially different ways. In the first place we can mix finite amounts of the two pure components, e.g., n_1 molecules of K_1 and n_2 molecules of K_2 : this is termed an *integral mixing process*, and the maximum work obtainable will be denoted by A . In the second place, we can dissolve a certain amount of a pure component, e.g., one molecule, in an infinite quantity of a mixture of K_1 and K_2 , so that the concentra-

tion of the mixture, $x = \frac{n_1}{n_1 + n_2}$ is not altered. This is termed a

differential mixing process. If it is carried out with one molecule of either component, the differential work of mixing obtained will be denoted by A_1 and A_2 respectively. For mixing processes carried out at constant temperature and pressure there exist important relations between the three quantities A_1 , A_2 and A .

An isothermal integral mixing process giving work A can be carried out differentially by simultaneously adding $n_1 K_1$ and $n_2 K_2$ to a mixture of composition x in such a way that the concentration remains sensibly constant throughout. Since the work obtainable by the two methods must be the same, we have

$$A = n_1 A_1 + n_2 A_2. \quad (14)$$

The integral mixing process can be carried out by taking a given amount $n_2 K_2$ of the second component and adding to it $n_1 K_1$ in infinitesimal amounts dn_1 , so that the addition of each amount is a differential process. We then have obviously

$$dA = A_1 dn_1, \quad (15)$$

or, more precisely,

$$A_1 = \left(\frac{dA}{dn_1} \right)_{T, p, n_2}. \quad (16)$$

Analogously, we have for the other component

$$A_2 = \left(\frac{dA}{dn_2} \right)_{T, p, n_1}. \quad (17)$$

Since A is a function of n_1 and n_2 , the general rule for partial differential coefficients gives

$$dA = \left(\frac{dA}{dn_1} \right)_{T, p, n_2} dn_1 + \left(\frac{dA}{dn_2} \right)_{T, p, n_1} dn_2, \quad (18)$$

which in conjunction with (16) and (17) leads to

$$dA = A_1 dn_1 + A_2 dn_2. \quad (19)$$

Finally, complete differentiation of (14) gives

$$dA = n_1 dA_1 + A_1 dn_1 + n_2 dA_2 + A_2 dn_2, \quad (20)$$

which in conjunction with (19) gives the important relation between the two differential works,

$$n_1 dA_1 + n_2 dA_2 = 0. \quad (21)$$

The method of derivation shows that all these equations are only valid for constant temperature and pressure.

In illustrating this relation graphically it is best to use the x concentration scale, and to deal with A_i , the integral work of mixing per molecule of mixture, instead of A , which refers to $n_1 + n_2$ molecules of mixture. The corresponding equations can be derived analogously to those given above, or can be obtained from them by the relations

$$A = (n_1 + n_2) A_i,$$

and

$$x = \frac{n_1}{n_1 + n_2}.$$

This leads to the following relations, valid for constant temperature and pressure,

$$A_i = x A_1 + (1 - x) A_2, \quad (22)$$

$$\frac{dA_i}{dx} = A_1 - A_2, \quad (23)$$

$$A_1 = A_i + (1 - x) \frac{dA_i}{dx}, \quad (24)$$

$$A_2 = A_i - x \frac{dA_i}{dx}, \quad (25)$$

together with the important relation between the differential work terms,

$$x \frac{dA_1}{dx} + (1 - x) \frac{dA_2}{dx} = 0. \quad (26)$$

The relation between A_1 , A_2 and A_i expressed by these equations is shown in Fig. 2. It may be specially noted that the A_1 and A_2 curves intersect at the maximum of the A_i curve, and that they

approach the ordinate axes asymptotically at $x = 0$ and $x = 1$ respectively.

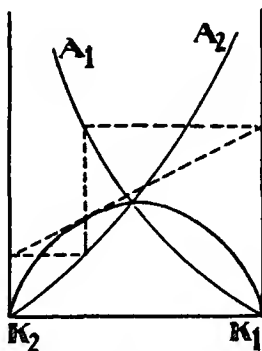


Fig. 2.

Equations (21) and (26) can also be deduced directly by the following simple argument. If one gram-molecule of mixture is transferred from concentration x to concentration $x + dx$, the work obtained must be infinitesimal since the concentration difference is infinitesimal. If further the composition of the gram-molecule transferred differs infinitesimally from the composition of the solutions between which it is transferred, the work obtained must be reduced to an infinitesimal of the second order, i.e., it must be effectively zero. The mixture $xK_1 + (1 - x)K_2$ has the required composition, and the work of transferring one molecule of this mixture must be zero, i.e.,

$$xdA_1 + (1 - x)dA_2 = 0.$$

(b) **General Relations between Differential and Integral Quantities.** It is easily seen that relations analogous to those derived above will be valid not only for the work of mixing, but also for any other extensive quantities associated with mixing which depend only on the quantities of the substances present. It is therefore necessary that the pressure and temperature should be kept constant during all the operations performed, since the properties of a system will in general depend on these two variables.

We can therefore write in general for the relation between an integral quantity X and the corresponding differential quantities X_1 and X_2 , T and p being constant,

$$X = n_1 X_1 + n_2 X_2, \quad (27)$$

$$X_1 = \left(\frac{\partial X}{\partial n_1} \right)_{T, p, n_2}, \quad (28)$$

$$X_2 = \left(\frac{\partial X}{\partial n_2} \right)_{T, p, n_1}. \quad (29)$$

$$n_1 dX_1 + n_2 dX_2 = 0 \quad (30)$$

Thus, for example, if these general equations are applied to the volume, $X = V$ is the integral or *total* volume of $(n_1 + n_2)$ molecules of mixture, while $X_1 = V_1$ and $X_2 = V_2$ are the *differential* molar solution volumes, *i.e.*, the increase of volume caused by dissolving one molecule of K_1 and K_2 respectively in an infinite amount of mixture.

In the same way equations (27) to (30) can be used to define *differential* thermodynamic functions X_1 and X_2 , which are of great importance for the thermodynamics of solutions. Thus the differential G -function defined by

$$\mu_1 = G_1 = \left(\frac{\partial G}{\partial n_1} \right)_{T, p, n_2}, \quad (31)$$

is identical with the chemical potential [I. 3. d.], and is equal to the increase of thermodynamic potential caused by adding one molecule of the component K_1 to an infinite amount of mixture of composition x . Similarly, the differential heat content,

$$H_1 = \left(\frac{\partial H}{\partial n_1} \right)_{T, p, n_2}, \quad (32)$$

is the increase of heat content when one molecule of K_1 is dissolved in an infinite amount of mixture. $H_1 - H_{1(1)}$ is easily seen to be identical with the differential heat of solution Q_1 .

It is important to realise that all the formulæ derived in Chapter I for the variations of the *total* thermodynamic functions can be transformed into equally valid formulæ by introducing differential functions in place of total ones. For example, equation I. (34)

$$\left(\frac{\partial G}{\partial p} \right)_T = V, \quad (33)$$

becomes

$$\left(\frac{\partial G_1}{\partial p} \right)_T = V_1. \quad (34)$$

To demonstrate this we shall first note that since n_1 and n_2 have been introduced as variables, equation (33) should be written

$$\left(\frac{\partial G}{\partial p}\right)_{T, n_1, n_2} = V. \quad (35)$$

Since, however, T and n_2 are kept constant in all the subsequent operations, we shall omit them for the sake of simplicity, and write

$$\left(\frac{\partial G}{\partial p}\right)_{n_1} = V. \quad (36)$$

Equation I. (44) now gives for constant T and n_2

$$dG = G_1 dn_1 + V dp, \quad (37)$$

and if we apply to this equation the differentiation rules in I. (72)–(74) we obtain

$$\left(\frac{\partial G_1}{\partial p}\right)_{n_1} = \left(\frac{\partial V}{\partial n_1}\right)_p. \quad (38)$$

However, we have from (28)

$$\left(\frac{\partial V}{\partial n_1}\right)_p = V_1, \quad (39)$$

and hence finally

$$\left(\frac{\partial G_1}{\partial p}\right)_{T, n_1, n_2} = V_1. \quad (40)$$

We thus obtain an equation for the differential quantities G_1 and V_1 which is quite analogous to the equation (35) valid for the total functions G and V . The effect of temperature upon G_1 is analogously given by

$$\left(\frac{\partial G_1}{\partial T}\right)_{p, n_1, n_2} = -S_1. \quad (41)$$

The proof assumes a similar form in all cases.

Just as equations (27) to (30) have been derived as the basis of the proof given for the special case $X = A$, so equations (22) to (26) can be transformed into general relations by the same substitution. We then obtain,

$$X_i = xX_1 + (1 - x)X_2, \quad (42)$$

$$X_1 = X_i + (1 - x)\left(\frac{\partial X_i}{\partial x}\right)_{T, p}, \quad (43)$$

$$X_2 = X_i - x\left(\frac{\partial X_i}{\partial x}\right)_{T, p}, \quad (44)$$

$$\left(\frac{\partial X_1}{\partial x}\right)_{T,p} + (1-x)\left(\frac{\partial X_2}{\partial x}\right)_{T,p} = 0 \quad (45)$$

where X_i is the value of the function X per unit amount of the mixture. These equations are useful in the graphical representation of the variation of the thermodynamic functions with concentration (cf. Figs 2 and 11). For example, it is easily seen that the value of X_1 is given by the intercept on the ordinate $x = 1$ made by the tangent to the X_i -curve at the concentration x .

Equation (45) shows that at $x = 0$ we must have either

$$\left(\frac{\partial X_1}{\partial x}\right)_{T,p} = \infty \quad \text{or} \quad \left(\frac{\partial X_2}{\partial x}\right)_{T,p} = 0.$$

The first alternative is obeyed by the work of mixing and the functions F and G , while the second is obeyed by the functions E and H and the volume for mixtures of non-electrolytes.

(c) **The Chemical Potential. The Gibbs-Duhem Equation.** The chemical potential is defined by equations I. (40), (42), (44) and (45). We shall use here the definition in terms of the function G , which is specially simple,

$$dG = -SdT + vdp + \sum \mu_i dn_i,$$

$$\mu_1 = \left(\frac{\partial G}{\partial n_1}\right)_{T,p,n}, \quad \mu_2 = \left(\frac{\partial G}{\partial n_2}\right)_{T,p,n} \dots$$

These equations show (as does equation (31)) that μ is identical with the differential G -function:

$$\mu_1 = G_1, \quad \mu_2 = G_2 \dots, \quad (46)$$

so that we can write from (27) and (30),

$$G = n_1\mu_1 + n_2\mu_2, \quad (47)$$

$$n_1d\mu_1 + n_2d\mu_2 = 0. \quad (48)$$

This equation can also be obtained directly from I. (56) for T and p constant. It is easily seen that an analogous equation holds for any number of components. We thus have in general

$$\sum n_i d\mu_i = 0, \quad (49)$$

the conditions for the validity of both equations being that T and p are constant.

(48) and (49) represent the so-called Gibbs-Duhem equation, which is of great importance for the thermodynamics of mixtures. If one of the components, e.g., K_1 , is present at infinitesimal concentration, the corresponding differential coefficient $\left(\frac{\partial \mu_1}{\partial n_1}\right)_n$

will be infinite compared with $\left(\frac{\partial\mu_2}{\partial n_1}\right)_n$. This follows from V. (3) and (4).

I. (52) shows that

$$\mu_1 - \mu_{1(1)} = -A_1 + p(V_1 - V_{1(1)}). \quad (50)$$

For liquid systems the last term in this expression is in general of little importance, and we can therefore write as a good approximation for such systems,

$$\mu_1 - \mu_{1(1)} = -A_1, \quad \mu_2 - \mu_{2(2)} = -A_2 \dots, \quad (51)$$

i.e., we can identify the increase of chemical potential associated with the mixing process with the negative value of the differential work of solution.

Since the potential is identical with the differential G -function, (40) and (41) give directly

$$\left(\frac{\partial\mu_1}{\partial p}\right)_{T,x} = V_1 \quad (52)$$

and

$$\left(\frac{\partial\mu_1}{\partial T}\right)_{p,x} = -S_1. \quad (53)$$

By introducing (52) and (53) in the general differential equation

$$d\mu_1 = \left(\frac{\partial\mu_1}{\partial T}\right)_{p,x} dT + \left(\frac{\partial\mu_1}{\partial p}\right)_{T,x} dp + \left(\frac{\partial\mu_1}{\partial x}\right)_{T,p} dx, \quad (54)$$

we obtain

$$d\mu_1 = -S_1 dT + V_1 dp + \left(\frac{\partial\mu_1}{\partial x}\right)_{T,p} dx, \quad (55)$$

or, for constant composition,

$$d\mu_1 = -S_1 dT + V_1 dp. \quad (56)$$

(d) **Ideal Mixtures.** It is not possible in general to predict how the chemical potential of a component will depend on the composition of the solution, since the nature of the dependence varies with the chemical nature of the components, just as for the work of mixing. A knowledge of the potentials as functions of the concentration gives a complete thermodynamic description of the mixture in question at constant temperature and pressure.

Certain liquid mixtures are characterised by particularly simple behaviour, the chemical potentials over the whole concentration

range being found to obey the same laws as those deduced for gaseous mixtures. For such liquid mixtures we have from equation (10),

$$\left. \begin{aligned} \mu_1 - \mu_{1(1)} &= RT \ln x, \\ \mu_2 - \mu_{2(2)} &= RT \ln (1 - x). \end{aligned} \right\} \quad (57)$$

These equations can of course only be tested by introducing definite values for the molecular weights M_1 and M_2 of the two components, since these quantities are needed to determine both the molecular potentials and the composition x . Equation (57) can therefore also be regarded as a definition of the molecular weight in the liquid state for ideal mixtures, and hence also for the molecular weight of the two components in the pure liquid state. It is found in general that the molecular weight in ideal liquid mixtures is the same as that of the pure components in the vapour state.

Mixtures which obey equation (57) will also behave ideally in other respects. Just as for gas mixtures, the change of energy and the heat of mixing at constant pressure and temperature must be zero, and the differential volumes will be independent of the concentration. We thus have for such mixtures

$$\left. \begin{aligned} V_1 &= V_{1(1)}, \\ V_2 &= V_{2(2)}. \end{aligned} \right\}$$

Equation (51) will therefore be not only approximately, but exactly valid, so that we can write

$$\left. \begin{aligned} A_1 &= -RT \ln x, \\ A_2 &= -RT \ln (1 - x). \end{aligned} \right\} \quad (58)$$

On the other hand, the equation $\left(\frac{dE}{dv}\right)_T = 0$, which is characteristic of perfect gases, will not be obeyed by ideal liquid mixtures any more than it is by pure liquids.

Mixtures which behave ideally are usually formed from components having molecules of similar structure. Thus mixtures of ordinary water with heavy water (deuterium oxide), or of toluene with benzene or diphenylmethane, obey the ideal laws approximately. The work of mixing can then be calculated from the equations for ideal gases. Thus the work obtainable by mixing $\frac{1}{2}$ gram-molecule of benzene and $\frac{1}{2}$ gram-molecule of toluene is given by equation (2) as

$$A_1 = -RT \ln \frac{1}{2} = 404 \text{ cal.} = 172 \text{ kg. metres.}$$

As we shall see in the next chapter, the laws of ideal solutions will hold for all solutions in the neighbourhood of $x = 0$ and $x = 1$, i.e., for what are termed "dilute solutions."

(e) **Activities and Activity Coefficients.** In describing the thermodynamic behaviour of mixtures, the chemical potential μ can be replaced by the activity a . The two functions are connected by the relation,

$$\mu_1 = RT \ln a_1 + k_1, \quad (59)$$

$$\mu_2 = RT \ln a_2 + k_2, \quad (60)$$

where for given components k_1 and k_2 are independent of the concentration, but dependent on the pressure and temperature. Thus if k_1 and k_2 are regarded as constants, (59) and (60) are only valid for constant temperature and pressure.

The presence of the terms k_1 and k_2 (depending on the temperature and pressure) in the definition of the activity allows freedom in the conventional choice of an absolute value for the activity under varying external conditions. Thus in many cases it is convenient to put $a_2 = 1$ for $x = 0$, and $a_1 = 1$ for $x = 1$, independent of the temperature and pressure. The activity is then unity for each component in the pure state.

With this convention the equations defining the activity can be written

$$\left. \begin{aligned} \mu_1 - \mu_{1(1)} &= RT \ln a_1, \\ \mu_2 - \mu_{2(2)} &= RT \ln a_2, \end{aligned} \right\} \quad (61)$$

where $\mu_{1(1)}$ and $\mu_{2(2)}$ are the chemical potentials of the pure components at the same temperature and pressure as the mixture under consideration. It will be seen that the addition of a constant term to the potential is equivalent to multiplying the activity by a constant factor. (61) is generally valid if $\mu_{1(1)}$ is defined as the value of μ_1 when $a_1 = 1$. If $\mu_{1(1)}$ is defined as the potential for $x = 1$, the validity of (61) also depends on the convention that $a_1 = 1$ for $x = 1$.

Equations (57), (59) and (60) show that for ideal solutions the activity is proportional to the concentration. If the above convention is introduced, they become identical, i.e.,

$$\left. \begin{aligned} a_1 &= x, \\ a_2 &= 1 - x. \end{aligned} \right\} \quad (62)$$

For non-ideal solutions we can write

$$\left. \begin{aligned} a_1 &= x f_1, \\ a_2 &= (1 - x) f_2, \end{aligned} \right\} \quad (63)$$

where f_1 and f_2 are factors depending on the concentration which are termed *activity coefficients*. If these coefficients are introduced into (61) it becomes

$$\left. \begin{aligned} \mu_1 - \mu_{1(1)} &= RT \ln x f_1, \\ \mu_2 - \mu_{2(2)} &= RT \ln (1 - x) f_2. \end{aligned} \right\} \quad (64)$$

The activity coefficients can be regarded as a measure of the deviation of the mixture in question from ideal behaviour.

The effect of temperature and pressure upon activities and activity coefficients is derived from (52) and (53) by introducing (61) and (63). For the effect of pressure we have

$$\left(\frac{\partial (\mu_1 - \mu_{1(1)})}{\partial p} \right)_{T,x} = V_1 - V_{1(1)},$$

or

$$RT \left(\frac{\partial \ln a_1}{\partial p} \right)_{T,x} = RT \left(\frac{\partial \ln f_1}{\partial p} \right)_{T,x} = V_1 - V_{1(1)}, \quad (65)$$

showing that the activity and the activity coefficient will be independent of pressure if the two components mix without any change of volume.

The effect of temperature is given by

$$\left(\frac{\partial (\mu_1 - \mu_{1(1)})}{\partial T} \right)_{p,x} = -(S_1 - S_{1(1)}),$$

or

$$RT \ln f_1 x + RT^2 \left(\frac{\partial \ln f_1}{\partial T} \right)_{p,x} = -T(S_1 - S_{1(1)}),$$

or, introducing (46) and I. (30),

$$RT^2 \left(\frac{\partial \ln a_1}{\partial T} \right)_{p,x} = RT^2 \left(\frac{\partial \ln f_1}{\partial T} \right)_{p,x} = -(H_1 - H_{1(1)}) = -Q. \quad (66)$$

The activity and activity coefficient is thus independent of the temperature if the heat of mixing is zero.

By applying the Gibbs-Duhem equation to (48) in the forms (59), (60) and (64), it is seen that both the activities and the activity coefficients are connected by relations analogous to (48), *i.e.*,

$$n_1 d \ln a_1 + n_2 d \ln a_2 = 0, \quad (67)$$

$$n_1 d \ln f_1 + n_2 d \ln f_2 = 0. \quad (68)$$

Analogous expressions hold for systems with more than two components.

The concepts of activity and activity coefficients have proved of

particular value in treating dilute solutions, which will be dealt with in the next chapter.

Besides the activity function described above, the so-called "absolute activity" is of some importance. This is defined by the equation

$$\mu_1 = RT \ln a_1 + k_1, \quad (69)$$

which is analogous to (59), but differs from it in that k_1 in (69) is a constant depending on the temperature and pressure, but independent of the nature of the other components present. In the same way, the "absolute activity coefficient" is defined in terms of the absolute activity by the equation

$$a_1 = f_1 x_1. \quad (70)$$

It will be seen from these definitions that equilibrium between any two phases at the same temperature and pressure demands that

$$a' = a'', \quad (71)$$

and hence that

$$\frac{f''}{f'} = \frac{x'}{x''}. \quad (72)$$

(f) **Equilibrium between Liquid and Vapour.** Just as in the case of a pure liquid, a liquid mixture at a given temperature is characterised by a definite vapour pressure. The composition of the vapour will also be fixed if the composition of the liquid is given. The composition of the vapour will however in general differ from the composition of the liquid with which it is in equilibrium.

If the vapour pressure is small and the vapour thus dilute, we can apply the partial pressure law (1), giving

$$\left. \begin{aligned} p_1 &= x'p, \\ p_2 &= (1 - x')p, \end{aligned} \right\} \quad (73)$$

where x' is the mole fraction of K_1 in the vapour. The total vapour pressure p is thus equal to the sum of the partial pressures. If the vapour is so concentrated that (73) is not valid, we can use the

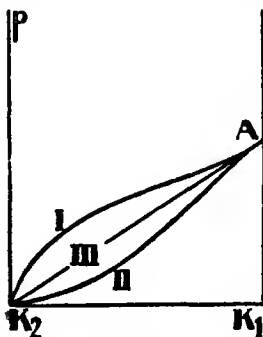


Fig. 3.

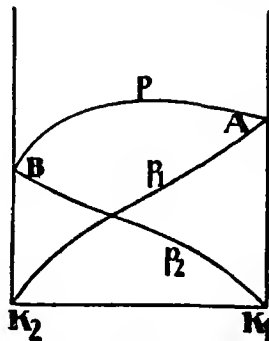


Fig. 4.

equation to define "theoretical partial pressures," which will of course by definition follow this law of additivity. The determination of p_1 and p_2 in practice usually involves (73), since it is based on measurements of x' and the total pressure p , or on measurements of the volume concentrations in the vapour.

Fig. 3 shows examples of the relation between p_1 , the partial pressure of the component K_1 , and x , its concentration in the liquid. The precise form of this curve cannot be predicted. However, it must of course begin at the point K_2 where $p_1 = 0$, and end at the point A , where $p_1 = p_{1(1)}$, the vapour pressure of pure K_1 . Further, as we shall see in the next chapter, the p_1 -curve at A must point towards the point K_2 . Finally, it can be established that for completely miscible components (e.g., alcohol and water), the p_1 -curve will rise continuously with increasing values of x . Exactly analogous laws will of course hold for p_2 .

Fig. 4 shows an example of the curves for both the two partial pressures and the total pressure, the abscissæ again being the values of x . The three curves are not mutually independent. To begin with, we must have $p = p_1 + p_2$, from (73). Further, in addition to the fixed nature of their terminal points, the shapes of the two partial pressure curves are not independent of one another.

The relation between them can be obtained if we remember that at equilibrium μ has the same value in the liquid and in the vapour. We can therefore write from (12),

$$\left. \begin{aligned} \mu_1 &= RT \ln p_1 + i_1, \\ \mu_2 &= RT \ln p_2 + i_2, \end{aligned} \right\} \quad (74)$$

as the relation between the chemical potentials in the liquid and the vapour pressures. If we neglect the small effect of the varying vapour pressure upon the potentials in the liquid (which is usually justifiable), we can apply the Gibbs-Duhem relation (48), giving

$$n_1 d \ln p_1 + n_2 d \ln p_2 = 0, \quad (75)$$

as the required relation between the partial pressures. This equation can also be written

$$\frac{x}{p_1} \frac{dp_1}{dx} + \frac{1-x}{p_2} \frac{dp_2}{dx} = 0. \quad (76)$$

By comparing (75) with (67), or more directly (59) with (74),

$$\begin{aligned} \mu_1 &= RT \ln a_1 + k_1, \\ \mu_1 &= RT \ln p_1 + i_1, \end{aligned}$$

we see that the activity and the vapour pressure are directly proportional. The vapour pressure of a component in a liquid mixture

thus provides a direct measure of the activity. This can be expressed by writing

$$a_1 = kp_1, \quad (77)$$

where the proportionality factor k is so far undetermined. Equation (77) shows that the partial pressure curves in Fig. 4 also express the variation of the activities with concentration.

It is clear that if the whole of the p_1 -curve is known, the shape of the p_2 -curve can be calculated from equation (75). If in addition a single point is known on the p_2 -curve or the p -curve (*e.g.*, the vapour pressure of the pure component $K_2, p_{2(2)}$), then the whole of the vapour pressure diagram can be constructed.

In the case of ideal mixtures, combination of (57) and (74) gives

$$\left. \begin{aligned} p_1 &= p_{1(1)}x, \\ p_2 &= p_{2(2)}(1-x), \end{aligned} \right\} \quad (78)$$

showing that in ideal mixtures the partial pressures are proportional to the molecular fractions of the components. The same must hence be true of the activities. In this case we therefore have *linear* vapour pressure curves, both for the partial pressures and the total pressure. If the actual curves lie above the linear ones, as curve *I* in Fig. 3, they are termed *positive*, while if they lie below they are termed *negative* vapour pressure curves. In certain cases there can be a maximum or a minimum in the total pressure curves, but, as already stated, these cannot occur in the partial vapour pressure curves of completely miscible liquids.

From the point of view of molecular theory, the different types of vapour pressure curves can be roughly characterised in the following way. Positive vapour pressure curves occur in mixtures where the molecular attractions K_1-K_1 and K_2-K_2 are stronger than the molecular attraction K_1-K_2 . If the reverse is the case, negative curves are obtained. Linear curves will therefore be restricted to the case in which the intermolecular forces between all the molecules are the same.

(g) **Vaporisation and Distillation of Mixtures.** The change in the composition of a liquid on vaporisation or distillation will of course depend on the relation between the composition of the liquid and the composition of the vapour which is removed. Even in rapid vaporisation, the vapour will usually be in equilibrium with the liquid during the process.

Information about these changes can be obtained from the total vapour pressure curves if we represent in the diagram not only the composition of the *liquid*, x , but also the composition of the *vapour*, x' . The diagram will then contain two curves, the liquid and vapour curves, or (in other words) the vapour pressure and dew-

point curves. The relative positions of these two curves obey the following general rules.

If there is equilibrium between liquid and vapour, then an isothermal vaporisation (caused by an increase in volume) will produce a decrease of vapour pressure. If the vaporisation caused an increase of vapour pressure, a certain amount of vaporisation will bring about further vaporisation, and the state would not be a stable one. It therefore follows that during evaporation the composition of the liquid will be changed in the direction of the component which causes a decrease of vapour pressure when it is added to the mixture, and that the vapour phase is relatively richer in that component which causes an increase in vapour pressure when

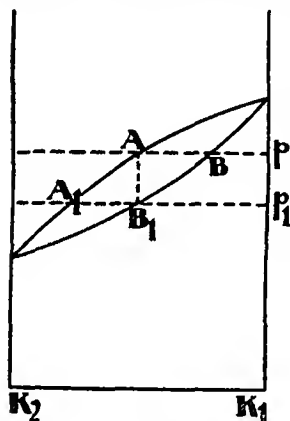


Fig. 5.

added to the mixture (Konowálow's rule). In Fig. 5, if A is a point on the liquid curve, then since the liquid curve rises from K_2 to K_1 , the corresponding point B on the vapour curve must lie to the right of A . The same must be true of the relative positions of the liquid and vapour curves over their whole course.

If one of the phases is present in infinitesimal amount, the composition of the mixture which is converted from liquid to vapour will be equal to the composition of this phase. If the ratio between the amounts of the phases is finite, the composition of the mixture evapo-

rated will be intermediate between the compositions of the two phases.

The relative positions of the liquid and vapour curves can also be deduced by considering the partial vapour pressure curves. The total vapour pressure is increased by adding that component which has the steepest partial pressure curve. Equation (76) then shows that the same component must be present in relatively higher concentration in the vapour phase, where the concentrations are proportional to the partial pressures.

Finally, Konowálow's rule can be deduced by the following thermodynamic argument, which also expresses the rule quantitatively and extends its application to any two-phase equilibrium.

If two phases denoted by ' and '' are in equilibrium at constant temperature, equation I. (56) gives

$$\left. \begin{aligned} -v'dp + n_1'd\mu_1 + n_2'd\mu_2 &= 0, \\ -v''dp + n_1''d\mu_1 + n_2''d\mu_2 &= 0, \end{aligned} \right\} \quad (79)$$

or

$$v' - v'' = (n_1' - n_1'') \left(\frac{\partial \mu_1}{\partial p} \right)_T + (n_2' - n_2'') \left(\frac{\partial \mu_2}{\partial p} \right)_T,$$

where the two differential coefficients, expressing the change of μ_1 and μ_2 in the equilibrium state per unit pressure change, must of course have opposite signs. Putting $n_1' + n_2' = n_1'' + n_2'' = 1$, and denoting by V' and V'' the volume of a gram-molecule of mixture in the two phases, we have

$$\left. \begin{aligned} V' - V'' &= (x' - x'') \left(\frac{\partial \mu_1}{\partial p} \right)_T - (x' - x'') \left(\frac{\partial \mu_2}{\partial p} \right)_T = \\ &= (x' - x'') \left[\left(\frac{\partial \mu_1}{\partial p} \right)_T - \left(\frac{\partial \mu_2}{\partial p} \right)_T \right]. \end{aligned} \right\} \quad (80)$$

If $V' > V''$ and $\left(\frac{\partial \mu_1}{\partial p} \right)_T$ is assumed positive, then $\left(\frac{\partial \mu_2}{\partial p} \right)_T$ is negative and the right-hand side of the equation has the same sign as $x' - x''$, and hence as $V' - V''$. The component, the addition of which increases the vapour pressure, will therefore be enriched in the phase which has the greater molar volume in the equilibrium state. Equation (80) gives a general quantitative expression for the effect of pressure on phase equilibria in binary systems.

Continuous isothermal vaporisation of the mixture *A* (Fig. 5) will first form an infinitesimal amount of vapour of composition *B*, after which the composition of the liquid will move from *A* to *A*₁ while that of the vapour will move from *B* to *B*₁, the last point corresponding to the same composition as *A*. At this point the liquid will all have evaporated and there will remain only an infinitesimal amount of liquid of composition *A*₁. This point therefore represents the limiting state for the liquid when vaporisation is carried out so that the liquid is always in equilibrium with the vapour formed.

If vaporisation is interrupted before it is complete, *e.g.*, at the pressure *p*₁ (Fig. 6), the original liquid of composition *A* will have been split up into the liquid (residue) *A*₁ and the vapour *B*₁, which after separation from the residue can be condensed to liquid of the

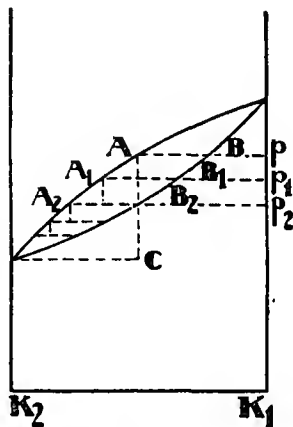


Fig. 6.

same composition. The liquid A_1 can now be treated in the same way and evaporated until the pressure is p_2 , when it will be separated into fractions of compositions A_2 and B_2 . By continuing this procedure we can obtain pure component K_2 , though only in very small quantity.

The ratio of the amounts of the phases A_1 and B_1 formed from the original liquid A is given in terms of the appropriate concentrations x_A , x_{A_1} and x_{B_1} , since for purely stoichiometrical reasons we have

$$\frac{A_1}{B_1} = \frac{x_{B_1} - x_A}{x_A - x_{A_1}},$$

independent of the nature of the phases.

If the vapour is condensed continuously as it is produced, the

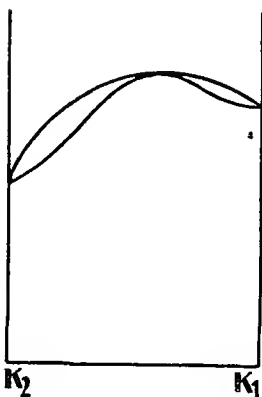


Fig. 7.

composition of the residue will change from A to pure component K_2 , while that of the condensed vapour will change from B to C . In continuous distillation the composition of the distillate will be represented by a curve joining B and C to the right of a curve joining B and C to the right of the vapour curve.

If the curve for the total vapour pressure has a *maximum* or a *minimum* point, then the application of Konowálow's rule leads directly to the result that the liquid and vapour curves must touch one another at this point (Fig. 7). On distilling mixtures having concentrations which differ from the concentration at this point, it is easily seen that the composition of both the

liquid and the vapour will be displaced in the direction of lower vapour pressures. If there is maximum, the pure components will be finally obtained in infinitesimal quantities, while if there is a minimum, the mixture corresponding to the minimum will be obtained. For compositions corresponding exactly to the maximum or minimum point, *the liquid will vaporise and distil unchanged* like a pure substance. This also follows from (76) and (80), which show that the compositions of the liquid and vapour are identical at this point.

(h) **Vapour Pressure and Heat of Mixing.** If we apply equation II. (49),

$$q_1 = RT^2 \frac{d \ln p_1}{dT},$$

to the differential vaporisation of K_1 from a mixture, and then combine it with the corresponding equation for the pure component,

$$q_{1(l)} = RT^2 \frac{d \ln p_{1(l)}}{dT},$$

we obtain

$$q_1 - q_{1(l)} = -Q_1 = RT^2 \frac{d \ln \left(\frac{p_1}{p_{1(l)}} \right)}{dT}, \quad (81)$$

where Q_1 is the differential heat absorption on mixing, *i.e.*, ΔH_1 . This quantity depends on the temperature coefficient of the relative vapour pressure. For ideal solutions, $p_1/p_{1(l)}$ is independent of temperature, so that $Q_1 = Q_2 = Q_3 = 0$.

Equation (81) can also be derived direct from (66).

(i) **Mutual Miscibility. Liquid-liquid Equilibria.** Complete miscibility of two liquids depends either on chemical similarity between their molecules (in which case the mixtures approximate to

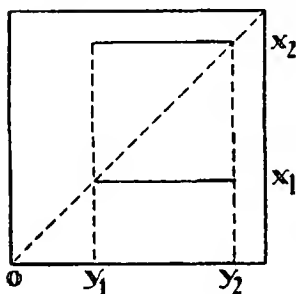


Fig. 8.

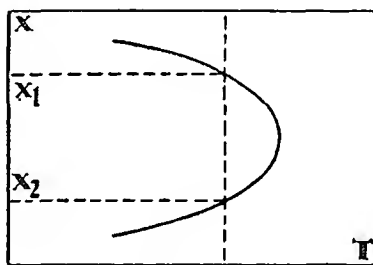


Fig. 9.

ideal behaviour), or on the presence of chemical affinity between the two sorts of molecules. The affinity which plays a part in the function of mixtures often depends on the *polarity* of the molecules, since polar molecules (*i.e.*, molecules in which the centres of positive and negative charge do not coincide [VII. 3. e.]) will become oriented in such a way that attraction results. On the other hand, a mixture of strongly polar and non-polar molecules will in general form an incompletely miscible system in which the equilibrium state consists of two liquid layers in contact. Molecules which are non-polar or only slightly polar usually behave most simply, and liquids consisting of such molecules are therefore termed "normal" liquids.

The system ether + water is an example of an incompletely miscible binary system. If we start with pure water, the addition

of a little ether gives a homogeneous mixture having a concentration x equal to the net composition of the mixture y . The same will be the case if a small quantity of water is added to pure ether. As long as the system is homogeneous, the relation between x and y will be represented on an x - y -graph as a straight line joining the points $x = y = 0$ and $x = y = 1$ (Fig. 8). If successive amounts of ether are added to water a net composition y_1 is eventually reached at which a new phase of composition x_2 separates out. If now further quantities of ether are added, the compositions of the two layers remain unchanged, but the amount of the x_2 -layer will increase at the expense of the x_1 -layer until we reach the net composition $y_2 = x_2$ at which the mixture is once more homogeneous. The homogeneous system now increases in ether concentration until $x = y = 1$.

An increase of temperature usually increases the mutual solubility of the two components, i.e., the two equilibrium concentrations x_1 and x_2 approach one another, as shown in Fig. 9, which represents the solubility-temperature curve. For example, phenol and water at 20° form a system of two liquid layers containing 8.4% and 72.2% of phenol. As the temperature increases, these concentrations approach one another, and at 68° they become identical, the two curves passing continuously one into the other. Above this so-called "*critical solution temperature*" the two liquids are completely miscible in all proportions. The equilibrium curve divides the diagram into two parts, points outside and inside the curve corresponding respectively to homogeneous and heterogeneous states.

There are, however, also examples of the reverse type of temperature dependence, and several cases of a so-called "*lower critical temperature*" have been realised. Thus the system triethylamine-water is completely miscible below 18°, but forms two layers in equilibrium above this temperature. Completely closed solubility curves are also known, e.g., the system nicotine-water, which has two critical solution temperatures, one at 60° and the other at 210°. When the point representing the state of the system passes the equilibrium curve two phases are formed; at the critical point these are equal in amount and differ infinitesimally in composition, while at any other point both the amounts and the compositions of the phases will be different.

Since the two layers are in equilibrium with one another, the chemical potential of each component must have the same value in the two phases. We therefore have, independent of the amounts of the two phases,

$$\mu_1' = \mu_1'', \quad (82)$$

where ' and '' denote the two phases having compositions x' and x'' respectively.

Identity of chemical potentials corresponds to identity of partial pressures. The partial pressure curves for incompletely miscible liquids must therefore always have both a maximum and a minimum. The two partial pressure curves illustrated in Fig. 10 show that this is most likely to happen for positive vapour pressure curves: in fact, systems with a high vapour pressure maximum will always be incompletely miscible.

The limit of miscibility can be expressed in terms of either chemical potentials or vapour pressures. The geometrical construction for X_1 described in connection with equations (42) to (45) shows that

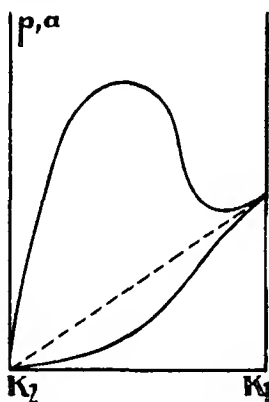


Fig. 10.

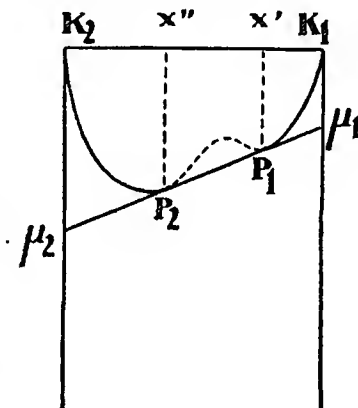


Fig. 11.

the value of μ_1 is completely determined by the tangent to the G_i -curve at the concentration corresponding to μ_1 . If (82) holds, the tangent to the G_i -curve at $x = x'$ must coincide with the tangent at $x = x''$. The condition for incomplete miscibility can therefore be expressed by the condition that a double tangent can be drawn to the G_i -curve. This is illustrated by the curve in Fig. 11, where it has been assumed for simplicity that $\mu_{1(1)} = \mu_{2(2)} = 0$. The points of contact P_1 and P_2 therefore represent the limits of miscibility. The G_i -function for the stable system is represented by $K_1P_1P_2K_2$, the portion between the points of contact, corresponding to heterogeneous states, being given by the straight line P_1P_2 . It follows from the figure that in stable mixtures the potential of a component must always increase with the concentration of that component, e.g., $d\mu_1/dx > 0$.

r.c.

1

The identity of the potentials and of the vapour pressures will of course extend to the compositions of the vapours in equilibrium with the two liquid phases. The vapour pressure-concentration curve must of course run parallel to the x -axis in the heterogeneous interval, as shown in Fig. 12.

(i) **Ternary Systems.** The composition of a phase containing three components can be represented by means of an equilateral triangle, the vertices of which correspond to the three pure components. The distance of a point in the triangle from one of the sides is a measure of the relative amount of the component corresponding

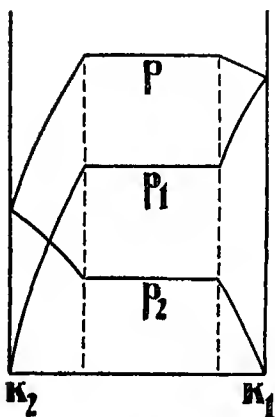


Fig. 12.

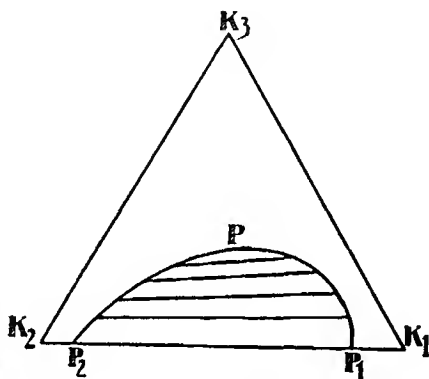


Fig. 13.

to the opposite angle. The three binary systems are represented by the three sides of the triangle.

If K_1 and K_2 are two incompletely miscible liquids, *e.g.*, ether and water, then the addition of a third substance, K_3 , can have either a positive or a negative effect upon the mutual solubility. For example, if the third substance is alcohol the miscibility will increase, until with continued additions a critical point will be reached beyond which there is complete miscibility. This is shown in Fig. 13, where P_1PP_2 is the solubility curve. Inside the curve are drawn chords, the ends of which correspond to *conjugate* solutions, *i.e.*, solutions which are in equilibrium with one another. At P , the critical point, the two ends of the chord coincide.

In other cases the addition of a third substance has the opposite effect. For example, if potassium carbonate is added to a mixture of water and alcohol, the mutual solubility will be decreased so that

the system separates into two liquid layers. This phenomenon is known as the *salting-out effect*.

(k) **The Solid-liquid Equilibrium.** If a two-component system consists of a solid and liquid phase in equilibrium with one another, then in most cases the liquid will be a homogeneous mixture and the solid either one of the pure components or a compound of both. In such an equilibrium system the solution is said to be *saturated* with respect to the solid phase. The *solubility* of the solid phase is given by the concentration of the saturated solution.

For a given system, the solubility will depend on the temperature and the pressure. We shall use the equilibrium conditions, expressed in terms of the chemical potential, to investigate this dependence.

At constant temperature the effect of pressure on the chemical potentials of component K_1 in the solid and dissolved states is given by (55) and I. (34) as

$$d\mu'_1 = V'_1 dp, \quad (83)$$

and

$$d\mu_1 = V_1 dp + \left(\frac{\partial \mu_1}{\partial n_1} \right)_p dn_1, \quad (84)$$

respectively, n_2 , the amount of K_2 being kept constant. In order that equilibrium shall be maintained for a change of pressure dp , it is necessary that these two changes in potential should be equal, giving as an expression for the effect of pressure

$$(V'_1 - V_1) dp = \left(\frac{\partial \mu_1}{\partial n_1} \right)_p dn_1, \quad (85)$$

or,

$$\frac{dn_1}{dp} = \frac{V'_1 - V_1}{\left(\frac{\partial \mu_1}{\partial n_1} \right)_p}, \quad (86)$$

or, emphasising the constancy of temperature,

$$\left(\frac{\partial x}{\partial p} \right)_T = \frac{V'_1 - V_1}{\left(\frac{\partial \mu_1}{\partial x} \right)_{T,p}}, \quad (87)$$

where x is the saturation concentration.

$\left(\frac{\partial x}{\partial p} \right)_T$ is thus the pressure coefficient of the solubility at constant temperature, and it can only be calculated quantitatively if we know

μ_1 as a function of the concentration at constant pressure and temperature. However, we know that for all stable solutions μ_1 increases with x . Equation (87) thus shows (in agreement with Le Chatelier's principle) that increase of pressure will increase the solubility if $V_1' > V_1$, i.e., if the solution of K_1 in the saturated solution is accompanied by a decrease of volume.

The effect of temperature on the solubility at constant pressure is derived analogously using (55) and I. (34),

$$d\mu_1' = -S_1' dT, \quad (88)$$

$$d\mu_1 = -S_1 dT + \left(\frac{\partial \mu_1}{\partial n_1} \right)_T dn_1. \quad (89)$$

Combining these equations, we have

$$(S_1 - S_1') dT = \left(\frac{\partial \mu_1}{\partial n_1} \right)_T dn_1, \quad (90)$$

or

$$\frac{dn_1}{dT} = \frac{S_1 - S_1'}{\left(\frac{\partial \mu_1}{\partial n_1} \right)_T}, \quad (91)$$

or, emphasising the constancy of pressure

$$\left(\frac{\partial x}{\partial T} \right)_p = \frac{S_1 - S_1'}{\left(\frac{\partial \mu_1}{\partial x} \right)_{T,p}}, \quad (92)$$

where x is the saturation concentration. The solubility will therefore increase with temperature if $S_1 > S_1'$, i.e., if the heat absorption when K_1 dissolves in saturated solution

$$q = (S_1 - S_1')T \quad (93)$$

is positive.

If the relation between μ_1 and x is known (as is the case for ideal mixtures and dilute solutions), then (87) and (92) can be used quantitatively for calculating the effect of pressure and temperature.

In many cases it is possible to realise a "complete solubility curve," i.e., an unbroken curve giving the variation of the solubility with temperature right up to the melting point of the pure solid phase, which can be considered as the natural end of the solubility curve. The general shape of such a curve is shown in Fig. 14. It is clear that the curve can equally well be regarded as a freezing point or melting point curve, i.e., a curve showing how the melting point of the component K_1 is altered by the addition of the component K_2 .

While points on the solubility curve correspond to saturation equilibrium, points above and to the left of the curve represent supersaturated states, and points below and to the right of the curve represent unsaturated states. A supersaturated solution is unstable, and will have a tendency to pass into a saturated solution with the deposition of the excess of the component in question. However, in the absence of the solid phase many supersaturated solutions exhibit considerable stability in practice and can be kept unchanged for long periods.

Fig. 15 represents the case in which the melting points of both components fall within the temperature interval investigated. These melting points are T_1 and T_2 . If the solubility curves have the same form as in the preceding example, they will intersect at a

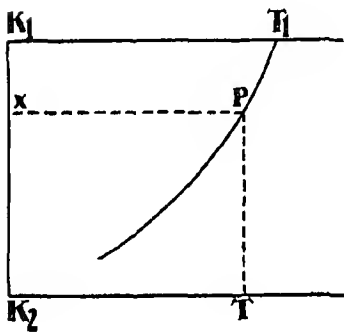


Fig. 14.

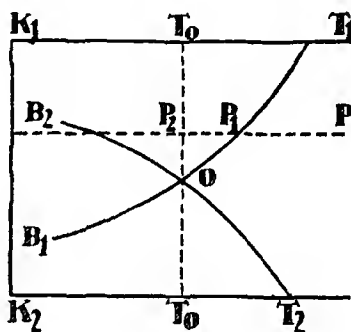


Fig. 15.

point O , corresponding to a concentration x_0 and a temperature T_0 . The point O is termed the *eutectic* point, or, in the case of solutions of salts in water, the *cryohydric* point: similarly, x_0 and T_0 are the eutectic (or cryohydric) concentration and temperature.

If a liquid mixture represented by the point P (concentration x) is cooled, it will remain homogeneous until the line representing its state cuts a solubility curve. This takes place at the point P_1 , and hence at this temperature the component K_1 will begin to separate out as a solid. If the cooling is continued, more K_1 will separate out, so that the liquid remaining in equilibrium with the solid phase will become richer in K_2 . The state of the liquid therefore moves along the solubility curve towards the eutectic point.

If it were possible to cross the solubility curve and reach the homogeneous point P_2 , this solution would be supersaturated with

respect to K_1 . The addition of a trace of solid K_1 would bring about crystallisation, and a point in the solubility curve would be reached. The same is true of any point to the left of the solubility curve for K_1 , B_1T_1 , which thus divides the diagram into two regions, the one on the left corresponding to supersaturated solutions, and the one on the right to unsaturated solutions.

Analogous considerations for K_2 show that solutions to the right of B_2T_2 are unsaturated, while solutions to the left of B_2T_2 are supersaturated with respect to K_2 . Solutions in the region B_1OB_2 are thus supersaturated with respect to both components.

If the state of the system passes the curve B_2T_2 from right to left it will be in a region where solid K_2 must be deposited in order to maintain equilibrium. If the change from P through P_1 and along the solubility curve B_1T_1 is continued, at the point O_2 , K_2 will begin to be deposited. Since the liquid is already in equilibrium with solid K_1 , the concentration of this component must not be altered by the deposition of K_2 : hence K_1 must also be deposited as a solid. Hence the whole liquid will solidify to a mixture of solid K_1 and solid K_2 without any change of concentration, *i.e.*, at the eutectic point the solution behaves like a pure substance in this respect. At temperatures below the eutectic temperature T_0 the stable state of all mixtures is a mixture of the two solid components.

In the diagram the area $K_1T_0T_0K_2$ represents this mixture of solids, the region T_0OT_1 represents solution + solid component K_1 , the region T_0OT_2 solution + solid component K_2 , and finally the region to the right of T_1OT_2 represents homogeneous solutions.

Since the region of stability for a mixture of solids lies below T_0 , two solids which are mixed at a higher temperature will melt partly or completely. This process absorbs heat, so that the temperature will sink. This behaviour is made use of in practice for producing low temperatures by means of *freezing mixtures*. The temperature lowering which can be produced depends on the form of the solubility curve and the heats of solution (or fusion) of the components. If the final temperature after mixing is T' and the composition of the mixture is x , the state of the system is represented by the point x, T' on the phase diagram. The point can lie in the region T_1OT_2 , in which case a homogeneous solution is obtained, or it can lie on one of the curves OT_1 and OT_2 , when a solubility equilibrium will be set up. It is easily seen that the temperature T_0 is the lowest temperature which can be produced in this way.

The table below gives the melting points and the eutectic temperatures and compositions for a number of substances of different kinds.

TABLE I. *Eutectic Points*

K_1	K_2	T_1-273	T_2-273	T_3-273	x
Silver . . .	Copper . . .	960°	1,081°	778°	60%
Bismuth . . .	Lead . . .	268°	322°	125°	56%
Acetic acid . . .	Water . . .	17°	0°	- 26°	31.1 mole %
Potassium iodide . . .	Water . . .	666°	0°	- 22°	5.3 "
Silver nitrate . . .	Water . . .	209°	0°	- 7.5°	8.3 "
NaCl . 2H ₂ O . . .	Water . . .	unknown	0°	- 22°	8.4 "
Potassium nitrate . . .	Sodium nitrate . . .	337°	308°	218°	46 "
Diphenylmethane . . .	Naphthalene . . .	27°	80°	14°	78 "

If the two components form a solid compound, the latter will of course have its own solubility curve. The solubility diagram can then assume various forms. Fig. 16 illustrates the case in which the melting point of the compound can be realised. The general form of the curve in this case can easily be constructed by imagining

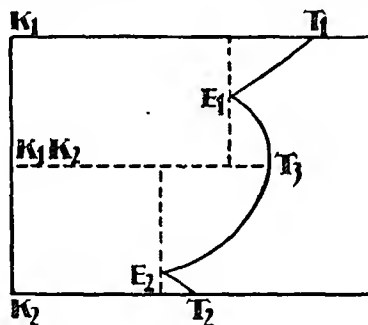


Fig. 16.

the system to be built up of two binary systems, $K_2 - K_1K_2$ and $K_1K_2 - K_1$, each of which behaves in the same way as the systems dealt with in the last paragraph. If we start with pure K_2 of melting point T_2 , the addition of K_1K_2 will lower the melting point along the curve T_2E_2 until it reaches the eutectic point E_2 , at which it meets the solubility curve of the compound. The solubility of the latter will increase with increasing temperature and the composition of the liquid will approach that of K_1K_2 , which is reached at the point T_3 , the melting point of the compound. At this point the solid compound gives on melting a liquid having the same composition as the solid. This phenomenon is known as *congruent fusion*. Addition of either K_1 or K_2 will lower the melting point, and T_3 is therefore a maximum on the melting point curve. If the compound

is partly dissociated into its components in the liquid, the addition of K_1 and K_2 will cause an abnormally small freezing point depression, and the melting point curve will be continuous with a vertical tangent at the point T_3 . The greater the extent of dissociation, the flatter the curve will be at the maximum. The portion of the curve $T_3E_1T_1$ is exactly analogous to $T_3E_2T_2$.

We may also show that the solubility curve of a compound is vertical at the point where the solid and liquid phases have the same composition by applying equation (92), which is equally applicable to the solubility of the pure components and of a compound between the components. If the solid phase is a compound of composition $\nu_1K_1 + \nu_2K_2$, a change in its chemical potential in solution is given by

$$\nu_1 d\mu_1 + \nu_2 d\mu_2 = d\mu,$$

while (48) gives in general for a homogeneous solution at constant temperature and pressure,

$$n_1 d\mu_1 + n_2 d\mu_2 = 0.$$

If $\frac{\nu_1}{\nu_2} = \frac{n_1}{n_2}$, i.e., if the solid and liquid phases have the same composition, this gives

$$\left(\frac{\partial \mu}{\partial x}\right)_{T,p} = 0,$$

and hence from (92),

$$\frac{dx}{dT} = \infty.$$

This means that the solubility curve of the compound has a vertical tangent.

In other cases, when the compound is heated in contact with saturated solution, it does not melt completely because it is converted to another compound or to a pure component before the melting point is reached. This type of behaviour is termed *incongruent* fusion, and is illustrated in Fig. 17. The solubility curve of the compound extends from E_2 to T_4 , at which point the stable solubility curve of K_1 begins. Hence at temperatures above T_4 the compound will break up into solution and K_1 , so that the portion of the curve T_4T_3 leading to the point of complete fusion T_3 cannot be realised as a stable state, since all its points obviously correspond to states which are supersaturated with respect to the component K_1 .

Thus when the compound is heated, partial fusion will take place when the point T_4 is reached.

This behaviour may be illustrated by an example. If water is one component and a salt the other, the compound is a so-called *salt hydrate*, e.g., $\text{Na}_2\text{SO}_4 \cdot 10\text{H}_2\text{O}$ or $\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$. The solubility diagram for the latter compound is given in Fig. 18. The lowest

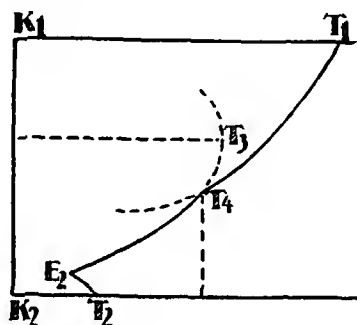


Fig. 17

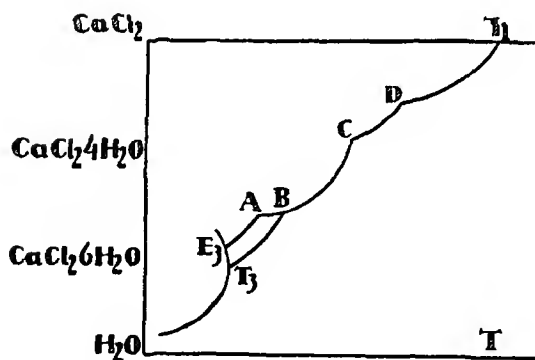


Fig. 18.

curve is the solubility curve of calcium chloride hexahydrate, $\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$, which extends from -50° (the cryohydric point) to 30.2° (T_3), the melting point of this compound. The curve for the hydrate $\beta\text{-CaCl}_2 \cdot 4\text{H}_2\text{O}$ runs from a point a little below T_3 to the point B. Another salt of the same composition, $\alpha\text{-CaCl}_2 \cdot 4\text{H}_2\text{O}$, has the solubility curve E_3A , but those parts of the curve which lie to the left of the solubility curve for the β -hydrate are unstable, since they correspond to supersaturated solutions. BC , CD and DT_1

are the solubility curves for the dihydrate, the monohydrate and the anhydrous salt, T_1 being the melting point of the latter compound. Every point at which two curves intersect corresponds to equilibrium between the solution and two solid salts. Thus at C there is equilibrium between monohydrate, dihydrate and solution; at E_3 between hexahydrate, α -tetrahydrate and solution, etc.

There can be essentially different types of equilibrium between two solid phases and a solution. An *eutectic point* corresponds to the intersection of two solubility curves when the composition of the solution lies between that of the two solid phases. The point of intersection is thus a temperature minimum for two stable solubility curves, the direction of which is otherwise undetermined. The other type of equilibrium corresponds to an intersection between two

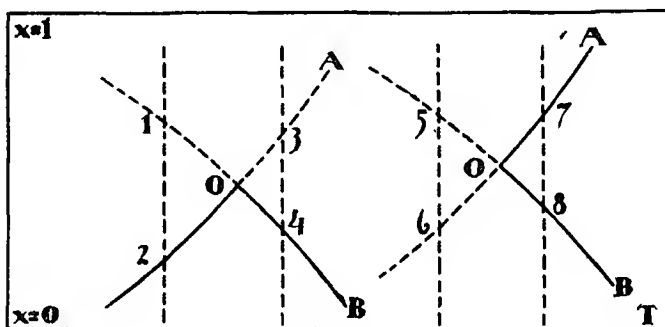


Fig. 19.

solubility curves when the composition of the solution is outside that of the two solids, so that the solution cannot be made from positive amounts of these two solids.

The different stability possibilities in a system where two solubility curves intersect are easily expressed by means of the following general rule. At a given temperature and pressure, if a solution *I* is in equilibrium with a solid phase, then a second solution *II* having a composition between that of the solid phase and that of the first solution will be unstable and will spontaneously split up into the two equilibrium phases. This follows because from a stoichiometric point of view solution *II* can be made by dissolving the solid phase in solution *I*: since solution *I* is already saturated, solution *II* must be supersaturated with respect to this solid phase.

This rule leads immediately to the result stated above about the stabilities and compositions of solid and liquid at an eutectic point and at a break in a solubility curve which is not a temperature

minimum. Fig. 19 shows two such points, the stable curves being represented by continuous lines and the unstable curves by broken lines. A and B represent the two solid phases to which the solubility curves refer. Since the point 1 is unstable, the above rule shows that the composition of A must lie above the point of intersection O , and since the point 3 is unstable, the composition of B must also lie above this point. The composition of the liquid thus lies outside those of the two solid phases. Further, since 5 is unstable, the composition of A must be above O , and since 6 is unstable, the composition of B must lie below O . Hence at an eutectic point the composition of the liquid lies between those of the two solid phases.

It is clear that the important factor in these considerations is not the angle between the curves, or their actual slopes, but is the positions of the unstable and stable portions of the curves relative to the axis of abscissæ. For example, the same results would be obtained if both curves rise with increasing temperature at the point O .

The melting point (or solubility) curves of binary mixtures have proved of particular importance in investigating salt hydrates and intermetallic compounds. They can, however, be used quite generally for detecting the existence of solid compounds between the components. Every temperature maximum corresponds to a chemical compound, the composition of which is determined by the composition of the liquid formed at the melting point; this can readily be determined in practice by analysis or in some other way.

We can thus state the following general law. If the compositions of two solid phases are represented by points lying above the composition of the point of intersection of their solubility curves, then the upper of the two curves corresponds to unstable states and the lower one to stable states. The break in the stable curve system will then point downwards, *i.e.*, increase of temperature causes a decrease in the gradient of the curve. If the compositions of the two crystals lie beneath the point of intersection, then the converse is of course true.

It should be noted that when dealing with a compound between the two components, the terms "greater or small solubility" have an indefinite meaning, since when two solubilities are being compared, an increase in the concentration of K_1 necessitates a decrease in the concentration of K_2 , and *vice versa*. We shall arbitrarily term that solubility the greatest which is nearest to the composition of the two solid phases.

(4) **The Vapour Pressure of Saturated Solutions.** If one component, *e.g.*, K_2 , is a volatile liquid while K_1 is a non-volatile solid which dissolves in the liquid, then by the stability rules the vapour

pressure must always decrease when x , the concentration of K_1 , increases. Hence if K_1 exists in two polymorphic forms, then a solution saturated with the stable form will have a greater vapour pressure than a solution saturated with the unstable form. It is easily seen that the same will be true if the solid phases contain a smaller proportion of K_2 than the saturated solutions.

A change of temperature has a dual effect upon the vapour pressure of a saturated solution, since on the one hand it affects the vapour pressure of a solution of constant concentration, while on the other hand it affects the solubility. The law for the variation of the vapour pressure with temperature can be derived in the following way:—

From I. (56) we have for the solution

$$SdT - Vdp + n_1d\mu_1 + n_2d\mu_2 = 0. \quad (94)$$

Further, for one gram-molecule of the vapour

$$d\mu_2' = -S_2'dT + v_2'dp, \quad (95)$$

and for one gram-molecule of the crystal,

$$d\mu_1'' = -S_1''dT + v_1''dp. \quad (96)$$

Introducing (95) and (96) into (94) and using the equilibrium conditions $d\mu_1 = d\mu_1''$ and $d\mu_2 = d\mu_2'$, we have

$$(S - n_1S_1'' - n_2S_2')dT - (V - n_1V_1'' - n_2V_2')dp = 0,$$

or, since we can neglect the volume of the solid and the solution compared with that of the vapour,

$$\frac{S}{n_2} - \frac{n_1}{n_2} S_1'' - S_2' = -V_2' \frac{dp}{dT}.$$

If we now introduce $pV_2'' = RT$, and $\Delta H = T\Delta S$ (valid for equilibrium), this equation becomes

$$q = RT^2 \frac{d \ln p}{dT}, \quad (97)$$

where

$$q = -\frac{H}{n_2} + \left(\frac{n_1}{n_2} H_1'' + H_2' \right), \quad (98)$$

is the heat absorbed in converting to solid and saturated vapour that quantity of solution which contains one gram-molecule of solvent. q is thus the molar heat of vaporisation of the solvent minus the integral heat of solution for the amount of dissolved substance needed to saturate one gram-molecule of solvent.

If l is the molar heat of vaporisation of the solvent and s the molar

integral heat of solution of the solute at saturation concentration, then (97) can be written in the form

$$q = l - s \frac{x}{1-x} = RT^2 \frac{d \ln p}{dT}.$$

As long as x is small, $\frac{dp}{dT}$ will be positive. If on the other hand x (i.e., the solubility of the solid) is large and s positive (which will always be the case near the melting point), $l - s \frac{x}{1-x}$ can be zero or negative. Hence if the solubility increases with increasing temperature the vapour pressure curve of the saturated solution may pass through a maximum and then fall again.

Curves of this kind are often met with for aqueous solutions of readily soluble salts. The form of the curve is shown in Fig. 20, where

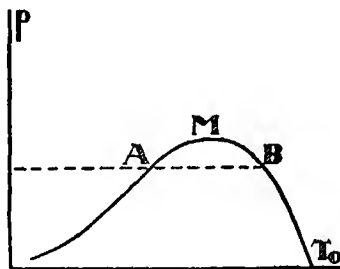


Fig. 20.

the curve is prolonged to T_0 , the melting point of the solute (assumed non-volatile). At this point $\frac{x}{1-x}$ is infinite, so that the curve for $\ln p$ must be vertical.

If the maximum M is at a pressure lower than 1 atmosphere, then the saturated solution obviously cannot be made to boil under ordinary atmospheric pressure (e.g., potassium hydroxide solution). If M lies above 1 atmosphere, then the saturated solution has two boiling points, A and B in the figure (e.g., sodium chlorate solution).

(m) **Equilibrium between One Solid and Two Liquid Phases.** In addition to the cases mentioned above, the complete course of a solubility curve can be interrupted by intersection with a curve representing the equilibrium between two liquid phases. This case is represented in Fig. 21. The curve D_1MD_2 is the equilibrium curve for the two liquid phases, and E is the eutectic point. The solubility curve for K_1 rises with increasing temperature from E to F_2 , where there must be equilibrium between the crystals of K_1 and the two liquid layers having compositions F_1 and F_2 . The

solubility curve for K_1 must then finish with the branch F_1T_1 leading to T_1 , the melting point of the pure component.

The temperature T corresponding to F_1 and F_2 is the melting point of K_1 "under the solvent," i.e., in presence of the solvent K_2 . Points to the left of the ordinate E_1E represent heterogeneous mixtures of the two solids K_1 and K_2 , the area $E_1EF_2F_1T_1$ equilibrium between K_1 and solution, the area $F_1MF_2F_1$ equilibrium between two liquid layers, and finally the whole region to the right of $T_1F_1MF_2ET_2$ to liquid homogeneous solutions. The curves D_1F_1 and D_2F_2 correspond to unstable equilibria. The type of

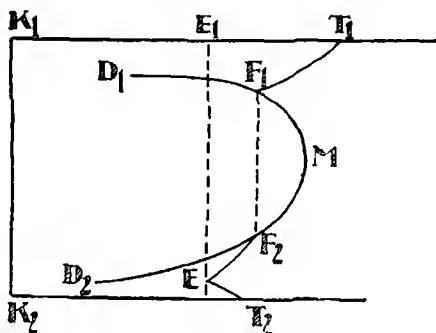


Fig. 21.

equilibrium corresponding to a given composition and a given temperature can be read directly from the diagram.

Diagrams of this kind are often met with, e.g., for phenol + water and benzoic acid + water. For the latter system M is at 114° , F_1 and F_2 at 90° , and T_1 at 121° .

4. CRYSTALLINE MIXTURES

When a crystal is formed by deposition from another phase, it will most often consist of a pure or almost pure substance. This is because miscibility in the crystalline state is in general very limited. However, for certain groups of substances the formation of homogeneous mixtures in the crystalline state is by no means unusual. A distinction has been made between two cases: in the first case the pure crystals of the two components are isomorphous and miscibility usually extends over the whole of the concentration interval, while in the second case there is no isomorphism and the miscibility is limited. The homogeneous crystals formed are termed respectively *mixed crystals* and *solid solutions* in the two cases.

There is, however, no fundamental distinction between the two groups of phenomena.

On account of the methods by which they were derived, the thermodynamic laws derived for liquid mixtures are also valid for mixtures which are crystalline. This applies both to the general relations derived in this chapter and to the special laws for dilute solutions treated in Chapter V. An important difference between solid and liquid mixtures lies in the much lower rate at which equilibrium is established in the former.

(a) **Vapour Pressures of Solid Solutions.** Only a few cases are known where a component of a solid solution is appreciably volatile. Certain silicates dissolve water to a small extent, and equilibrium can be established between the solid phase containing water and water vapour. This case is met with in particular in the so-called zeolites. In agreement with the equilibrium laws the vapour pressure increases continuously with increasing water concentration in the crystals and at low concentrations is directly proportional to the concentration.

Hydrogen dissolves in certain metals, *e.g.*, platinum and palladium, and over certain ranges of concentration behaves in the same way.

(b) **The Equilibrium Crystal Liquid.** The equilibrium between solid and liquid mixtures is analogous to that between liquid and gaseous mixtures. In the latter case it is most advantageous to consider the variation of pressure at constant temperature, while for the crystal-liquid equilibrium it is of greater interest to investigate the effect of temperature upon the equilibrium at constant pressure.

The composition of the solid and liquid phases in equilibrium with one another obey a rule analogous to Konowálow's rule for the liquid-vapour equilibrium, namely: *the solid phase contains relatively more of that component which causes an increase of melting point when added to the system.* This rule can be proved by considering an arbitrary equilibrium of this type between liquid and crystal. If under constant external conditions a small quantity of the solid phase melts, the result must be to diminish (or, in the limiting case, to leave unaltered) the tendency to melt, since if this tendency were increased the equilibrium could not be a stable one. A decrease in the tendency to melt must mean a rise in the melting point, since the heat of fusion is positive. The addition to the liquid of mixture having the same composition as the solid will thus increase the melting point, and the crystal must therefore be richer in the component which causes an increase of melting point.

This rule can also be deduced from equation I. (56), just like Konowálow's rule for the liquid-vapour equilibrium. Since we are

considering the behaviour at constant pressure, the equation can be written

$$\begin{aligned} S dT + n'_1 d\mu_1 + n'_2 d\mu_2 &= 0, \\ S' dT + n''_1 d\mu_1 + n''_2 d\mu_2 &= 0. \end{aligned}$$

By the same method as was used in deriving equation (80), this gives

$$-(S' - S) = (x' - x'') \left[\left(\frac{\partial \mu_1}{\partial T} \right)_p - \left(\frac{\partial \mu_2}{\partial T} \right)_p \right], \quad (99)$$

where S' and S'' are the entropies of one gram-molecule of mixture in the two states of aggregation. This equation gives a general quantitative expression for the effect of temperature on phase equilibria in binary systems.

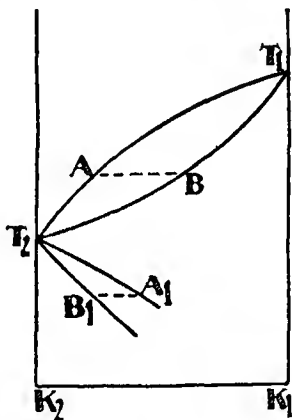


Fig. 22.

Fig. 22 illustrates the behaviour of this type of system. T_2AT_1 is the *liquidus curve*, representing the relation between the melting point and the composition of the liquid phase. Since addition of K_2 causes the melting point to fall, the liquid represented by the point A must contain a greater proportion of K_2 than the solid in equilibrium with it, represented by point B . B must hence lie to the right of A , and the *solidus curve* representing the relation between the melting point and the composition of the solid phase must lie to the right of the liquidus curve over the whole concentration range in which the latter rises with increasing content of K_1 .

In exactly the same way it is clear that solid B_1 corresponds to liquid A_1 , and that the liquidus and solidus curves must coincide if one of them has a horizontal tangent.

In the melting point diagram given in Fig. 23 the regions above $T_2A_2A_1T_1$ and below $T_2B_2B_1T_1$ will thus correspond respectively to homogeneous liquid and homogeneous solid systems, while points between the two curves correspond to heterogeneous states. A point on the line A_1B_1 will thus correspond to a mixture of liquid A_1 and solid B_1 . If a liquid corresponding to the point D is cooled, solidification will begin at A_1 , solid of composition B_1 being formed. On continued cooling the concentration of the liquid will be displaced from A_1 to A_2 , while that of the solid will pass from B_1 to B_2 . At the latter point the whole system will be solid, since the concentration at B_2 is equal to the concentration at A_1 . This description is

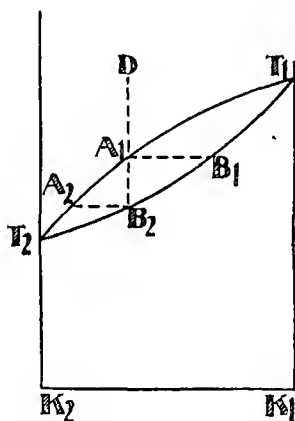


Fig. 23.

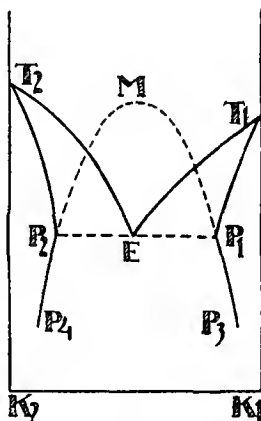


Fig. 24.

of course only correct on the assumption that there is always equilibrium between the solid and the liquid: this will only be true if the changes are carried out slowly or at a high temperature.

If the curves do not rise or fall continuously (as in Figs. 22 and 23), but pass through a maximum or a minimum melting point, this can occur in two essentially different ways. If the components are completely miscible both in the liquid and the solid state, then the composition of the phases in equilibrium with one another must vary continuously along the stable curves, and any maxima or minima occurring must lie on continuous curves. Fig. 7 can represent a diagram of this kind, the ordinates representing temperature.

If on the other hand the two liquidus curves intersect, as in Fig. 24, we have a minimum of a different type. At the point E , which is analogous to an ordinary eutectic point, the solid

phases P_1 and P_2 are both in equilibrium with the liquid E , and are hence in equilibrium with one another. If this is the case the solid phases cannot be completely miscible, and the points P_1 and P_2 representing the limits of mutual miscibility at the eutectic temperature must lie on a T - x -curve representing the variation of the mutual miscibility with temperature. Such a curve is analogous to the mutual miscibility curve for two liquid phases [3. m.], and can resemble it in having an upper critical point. This curve is represented by $P_4P_2MP_1P_3$, M being the critical point. However, all that part of the curve represented by a broken line is unstable, since it lies in regions corresponding to wholly or partially liquid states. In the figure the points above the line T_2ET_1 represent

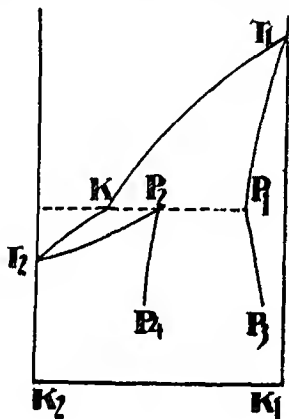


Fig. 25.

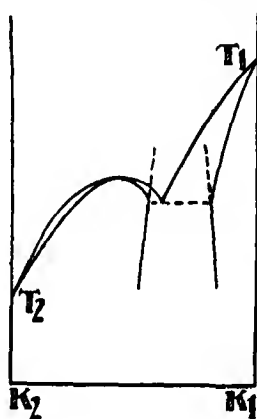


Fig. 26.

completely liquid homogeneous states. Outside $T_2P_2P_4$ and $T_1P_1P_3$ the system consists of solid mixed crystals containing predominantly K_2 and K_1 respectively, while in the areas T_2P_2E and T_1P_1E these two types of mixed crystals are in equilibrium with solution. Finally, below $P_4P_2P_1P_3$ the system consists of heterogeneous mixtures of the two types of homogeneous mixed crystals.

If the two liquidus curves intersect at a point intermediate between the two melting points a diagram resembling Fig. 25 is obtained, where as before the regions T_1P_1K , T_2P_2K and $P_3P_1P_2P_4$ represent heterogeneous states, the rest of the diagram representing homogeneous states.

Intersection of the two solidus curves in a stable region would only be possible if the liquid solutions were incompletely miscible and the solid solutions completely miscible: diagrams corresponding

to this case have not been realised in practice. It can be stated in general that a discontinuity in one of the equilibrium curves (*e.g.*, K in Fig. 25) corresponds to a gap in the series of solid solutions in the other equilibrium curve ($P_1 - P_2$).

If the curves pass through a maximum, then in this region there must either be complete miscibility both in the liquid and the solid phases, or there must be a chemical compound between the components. In the latter case the composition at the maximum will correspond to a simple molecular ratio between the components. Curves of this type are shown in Fig. 26.

5. SYSTEMS OF IMMISCIBLE COMPONENTS

(a) **Reactions between Solid Salts.** Reactions of this type, *e.g.*,



or

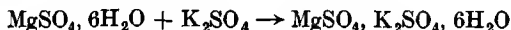


exhibit thermodynamic behaviour very similar to that of polymorphic transformations.

In general, however, the presence of a certain amount of solvent is necessary for the reaction to take place. The action of the solvent may be described as catalytic [VIII. 1. f.].

If one of the salts in the above salt systems is not present in solid form, its chemical potential in the mixture will nevertheless be completely determined by the temperature and the pressure [VI. 2. a.]. If this potential is smaller than that of the salt in the solid state, the system of solid salts will be stable. If on the other hand it is greater, a reaction such as the above will take place with deposition of the missing solid phase, thus stabilising the system.

At an arbitrary temperature and pressure, processes of this kind will take place completely either from left to right or in the reverse direction. At a given pressure there is only one temperature at which the two systems can be in equilibrium, corresponding to the "transition temperature" of polymorphic substances. The reason for this behaviour lies in the fact that the chemical potentials are independent of the extent to which the reaction has taken place. As is usual for processes in condensed systems, the effect of pressure is in practice small compared with that of temperature.

(b) **The Vapour Pressure of Salt Hydrates.** If equilibria of the above type also involve a vapour phase, the effect of pressure is of great importance. Important examples of this are provided by systems consisting of salt hydrates and water vapour. When a salt containing water of crystallisation gives up water to give a lower

hydrate (or an anhydrous salt), then if the initial and final systems are fixed a definite *equilibrium pressure* will be established at a given temperature. Thus, for example, sodium sulphate decahydrate at 20° will lose water until the vapour pressure has become about 14 mm., provided that the product of dehydration is the anhydrous salt (as is usually the case). If the partial pressure of water vapour is maintained at a lower value dehydration will take place completely, while conversely at a water vapour pressure greater than the equilibrium pressure the anhydrous salt will take up water and become completely converted to hydrate.

If a substance forms several hydrates the same principles apply,

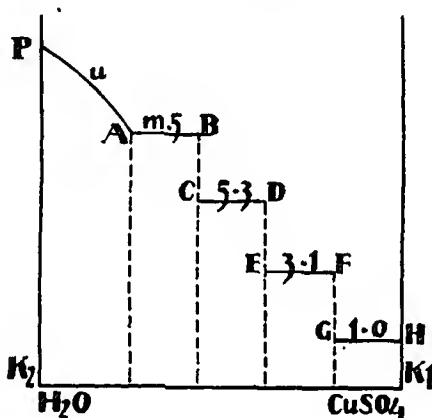


Fig. 27.

though the state of affairs is naturally more complicated. As an example we shall consider cupric sulphate, which forms hydrates with 5, 3 and 1 molecules of water, and for the sake of completeness we will investigate systems formed from water and cupric sulphate in all possible proportions.

In Fig. 27 the vapour pressure of pure water is represented by K_2P . The abscissæ represent the net composition of the solid and liquid phases taken together. On adding CuSO_4 the vapour pressure falls until it reaches the point *A*, where the liquid is saturated with the pentahydrate, $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$. If more CuSO_4 is added it will combine with water from the saturated solution to form pentahydrate, and the concentration of the solution will therefore remain unchanged. This is represented in the diagram by the horizontal line *AB*. At *B* all the solution has disappeared and the addition of

CuSO_4 or the removal of water will lead to formation of pentahydrate from the water vapour present until the pressure has fallen from *B* to *C*. If the proportion of CuSO_4 present is increased past the point *C*, trihydrate will begin to be formed according to the equation

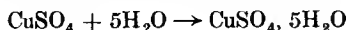


As long as both these hydrates are present the vapour pressure remains constant; at *D* the pentahydrate disappears and at *E* the monohydrate begins to be formed, etc. The three horizontal lines *CD*, *EF* and *GH* thus correspond to equilibria between two salts and vapour, while the vertical lines *BC*, *DE*, *FG* and *HK* correspond to the existence of a single salt, since the points contained in these lines represent the vapour pressures at which the single salts can exist. At a given temperature a pair of salts can thus co-exist only at a single point, while a single salt has a range of existence.

The same stages will of course occur if a dilute solution of cupric sulphate is gradually evaporated down. The existence of salt hydrates is most easily established experimentally by continuous removal of water in this way, the equilibrium vapour pressure being measured throughout.

If a salt hydrate is exposed to the atmosphere, it will lose water if its vapour pressure is greater than the partial pressure of water vapour in the atmosphere. This process is known as efflorescence. If this partial pressure is greater than the vapour pressure of a saturated solution of the salt, then water will be absorbed until all the salt has dissolved (deliquescence).

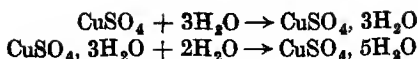
If it were possible to avoid the formation of $\text{CuSO}_4, 3\text{H}_2\text{O}$ and $\text{CuSO}_4, \text{H}_2\text{O}$ in the labile system $\text{CuSO}_4, 5\text{H}_2\text{O} + \text{CuSO}_4$, an equilibrium pressure will be developed which can be calculated in the following way. By means of equations previously derived [II. (52)], we see that the work obtainable from the process



is given by the expression

$$A = 5RT \ln \frac{p}{p_{5,0}},$$

where *p* is the vapour pressure of pure water and $p_{5,0}$ the equilibrium pressure in the labile mixture of pentahydrate + anhydrous salt. Since this process can also be written as the sum of the following two processes,



the work is also given by

$$A = 3RT \ln \frac{p}{p_{s,0}} + 2RT \ln \frac{p}{p_{5,3}},$$

where $p_{s,0}$ and $p_{5,3}$ are the vapour pressures in the systems trihydrate + anhydrous salt and pentahydrate + trihydrate; we thus have

$$5RT \ln \frac{p}{p_{5,0}} = 3RT \ln \frac{p}{p_{s,0}} + 2RT \ln \frac{p}{p_{5,3}},$$

whence

$$\frac{p_{s,0}^3 p_{5,3}^2}{p_{5,0}^5} = 1.$$

This shows that $p_{5,0}$ must lie between $p_{5,3}$ and $p_{s,0}$.

As the temperature rises, the vapour pressure of the hydrates must also rise, and we obtain the curves *a*, *b* and *c* in Fig. 28, giving

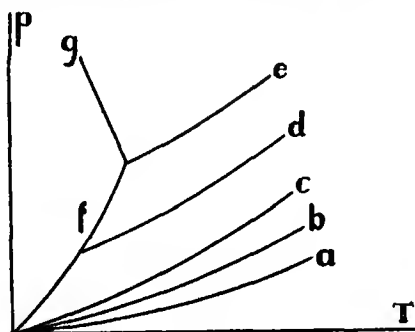


Fig. 28.

the variation of the equilibrium pressures with temperature. The curve *d* is the vapour pressure curve for solution saturated with pentahydrate, while *e* and *f* are the vapour pressure curves of water and ice respectively. The areas between the curves *d*, *c*, *b*, *a* and the temperature axis correspond to the existence of a single hydrate, while the curves correspond to the simultaneous existence of two hydrates, or of a hydrate and a saturated solution. The intersection of *d* and *f* constitutes a cryohydric point.

The intersection of the vapour pressure curves for two salt pairs merits special consideration. In the first place it is clear that the

three salts constituting the two salt-pairs can also form a third salt-pair. If the vapour pressure curves are given for two of the possible combinations, the curve for the third combination must also be determined. Further, if two of the curves intersect, the third curve must pass through the point of intersection. Thus, for example, if we assume that the curve for a pentahydrate-trihydrate system ($P-T$) intersects the curve for a trihydrate-anhydrous salt system ($T-A$) as shown in Fig. 29, then the curve for the pentahydrate-anhydrous salt system ($P-A$) must also pass through the point of intersection, since at this point all three salts are in equilibrium with the same vapour. It can be shown that outside the point

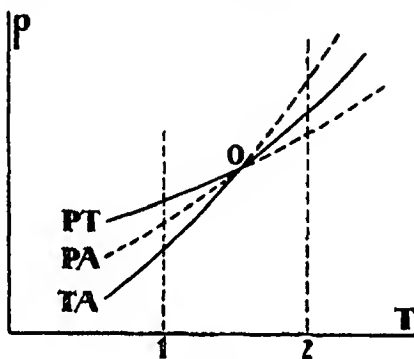


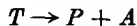
Fig. 29.

of intersection the $P-A$ -curve must lie between the $P-T$ and $T-A$ curves. This statement can be proved as above by considering the work obtained in the various processes, or it can be demonstrated in a more general way as follows.

Let us consider a temperature below the transition point and assume that at this temperature the three vapour pressures decrease in the order

$$p_{P-A} > p_{P-T} > p_{T-A}.$$

On account of the first of these inequalities the process



can take place spontaneously by the distillation of water vapour from $P-A$ to $P-T$, while on account of the second inequality the process



can take place spontaneously by distillation of water vapour from $P-T$ to $T-A$. But this contravenes the second law, since the two

processes differ only in direction, and if both are spontaneous they could be combined to produce a perpetual motion machine of the second kind: hence the original assumption about the vapour pressures must have been wrong.

Exactly the same result is obtained by considering the case

$$p_{P.T} > p_{T.A} > p_{P.A},$$

leaving as the only possibility the case in which the $T.A$ -curve lies between the two other curves.

These considerations tell us nothing about the stability relations, which are determined by the relative positions of the curves. Outside the transition point there are two possibilities, *i.e.*,

$$(1) \quad p_{P.T} > p_{T.A},$$

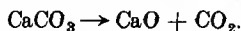
$$(2) \quad p_{T.A} > p_{P.T}.$$

If one of these inequalities holds, an arbitrary mixture of the three hydrates will not be stable. In case (1) the $P.T$ -system will lose water vapour, P being converted to T , while the $T.A$ system will absorb the water vapour liberated, A being converted to T . The process taking place is thus $P + A \rightarrow T$, and it will obviously continue as long as both P and A are present. This salt-pair is thus *unstable*. The process stops when either P or A is used up. In the first case $T.A$ is left, and in the second case $P.T$. Each of these salt-pairs is therefore *stable* separately.

If $p_{T.A} > p_{P.T}$, water vapour will be transferred in the reverse direction and the opposite processes will take place. $P.A$ is therefore *stable*, and $P.T$ and $T.A$ *unstable*. The instability of the last two pairs depends on the fact that T can both lose water (forming A) and take up water (forming P): hence T itself is unstable, being spontaneously converted into a higher and a lower hydrate. The net process taking place is thus the same as before, but in the reverse direction, $T \rightarrow P + A$. Case (1) can pass over into case (2) as the temperature changes. At the transition point $p_{P.T} = p_{P.A} = p_{T.A}$, and all three hydrates are stable simultaneously in the presence of saturated vapour.

Fig. 29 illustrates the case in which the first possibility holds at low temperatures and the second possibility at high temperatures.

The dissociation of a salt hydrate to form water vapour is formally quite analogous to more "chemical" processes, such as the dissociation of carbonates to form carbon dioxide, *e.g.*,



The laws governing dissociations of this type are therefore analogous to those derived above. They will be treated in a later chapter as a special case of homogeneous equilibria.

6. THE PHASE RULE

In considering the phase diagram for a single substance [II. 4. i.] it was established that points in the areas between the equilibrium curves represent a single state of aggregation (solid, liquid or gas), that points on the curves represent two phases in equilibrium, and that the point of intersection of two curves represents three phases in equilibrium. In the last case the state is completely determined, and such a point of intersection is a characteristic constant for a pure substance. This fact can also be expressed by saying that when a single substance is present as three phases, there are no *degrees of freedom*. If two phases, *e.g.*, liquid and vapour, are in equilibrium, there is one degree of freedom, since for example the temperature can be varied arbitrarily within certain limits, equilibrium being reached when the pressure assumes a definite value corresponding to the temperature selected. If a substance is only present as a single phase, both the temperature and the pressure can be altered arbitrarily, and the system thus has two degrees of freedom. For a pure substance there is thus the following connection between the number of phases f and the number of degrees of freedom n :—

f	n	$f + n$
3	0	3
2	1	3
1	2	3

The sum of the number of phases and the number of degrees of freedom is thus 3 for a single substance.

It can be shown that this is a special case of a general law known as the *phase rule* (Gibbs), which can be written

$$f + n = k + 2, \quad (100)$$

where k is the smallest number of components from which the system can be constructed with arbitrary ratios between the amounts of the phases. For a single substance this relation becomes $f + n = 3$, as found above.

Equation (100) can be derived in the following way. In a system of f phases and k components, equilibrium at constant temperature and pressure is attained when the potential of each component is constant

in all phases. The equilibrium condition can thus be expressed by the following equations:—

$$\begin{array}{ccccccc} \mu_1(1) = \mu_1(2) = & \dots\dots\dots & \mu_1(f), \\ \mu_2(1) = \mu_2(2) = & \dots\dots\dots & \mu_2(f), \\ \vdots & & \vdots \\ \mu_k(1) = \mu_k(2) = & \dots\dots\dots & \mu_k(f), \end{array}$$

the number of which is obviously $k(f-1)$.

The number of independent concentrations in a phase is $k-1$, since the concentrations are expressed as the relative quantities of the components present. The total number of concentrations is thus $f(k-1)$.

The state of the system can be varied by varying the $f(k-1)$ concentrations, the temperature and the pressure. There are thus $f(k-1) + 2$ variables which must satisfy the $k(f-1)$ equations between the potentials (which are functions of the concentrations, temperature and pressure). We thus have

$$\begin{aligned} L &= \text{number of equations} = kf - k \\ V &= \text{number of variables} = kf - f + 2 \end{aligned}$$

and hence

$$V - L = k - f + 2. \quad (101)$$

$V - L$ is the number of variables not determined by the equations: this is by definition equal to the number of degrees of freedom, so that equation (101) is identical to (100).

The above derivation shows that the phase rule is of general validity, and it is of great importance for the characterisation and classification of heterogeneous equilibria.

If an equilibrium has 0, 1, 2, 3, etc., degrees of freedom, the equilibrium and the corresponding system are termed invariant, monovariant, divariant, trivariant, etc. Monovariant equilibrium is often called complete equilibrium and is characterised by the fact that if one variable (e.g., the temperature) is arbitrarily varied while another variable (e.g., the pressure) is kept constant, then one phase will disappear completely. In a divariant (or incomplete) equilibrium a small temperature change at constant pressure will lead to a change in another variable (the concentration in a phase) sufficient to compensate the effect of the temperature change, and equilibrium will be maintained.

Invariant equilibrium corresponds to the greatest possible number of phases, namely $f = 3, 4, 5$, etc., for $k = 1, 2, 3$, etc. In the phase diagram this equilibrium is represented by a point, which is known as a triple, quadruple, quintuple point, etc., according to the number of phases. A corresponding number of equilibrium curves intersect at this point.

Many physico-chemical investigations are carried out under atmospheric pressure, and as an approximation the variations of pressure can be neglected. (100) then becomes

$$f + n = k + 1. \quad (102)$$

If on the other hand the extent of a surface is of importance [II. 2. a.], the number of degrees of freedom will be increased by unity, and we have

$$f + n = k + 3. \quad (103)$$

CHAPTER V

THE THEORY OF DILUTE SOLUTIONS

It was mentioned in the preceding chapter that the general laws valid for homogeneous mixtures assume particularly simple forms when one component is present in very small amount relative to the other. Mixtures of this kind are termed *dilute solutions*. The theory of these solutions constitutes an important chapter in classical physical chemistry, and its historical development was based on the quantity known as the *osmotic pressure*. It can however be shown that the laws governing this type of solution can be related to more simple experimental facts, and also to elementary considerations of molecular theory. We shall base our derivation of the laws of dilute solutions on a consideration of the expressions for potentials and activities in such solutions.

The expressions for mixtures dealt with in Chapter IV were completely symmetrical with respect to the two components. This symmetry will of course be absent in considering dilute solutions, where one component, the *solvent*, is present in very great concentration relative to the other, the *solute*. In order to emphasise this lack of symmetry we shall use a slightly different terminology in the following treatment: in particular we shall use the index 0 for the component present at high concentration, and the indices 1, 2, 3, etc., for the other components.

1. POTENTIALS AND ACTIVITIES

Dilute solutions are characterised by the fact that with increasing dilution their thermodynamic properties approach asymptotically to the properties of an ideal solution of the same concentration. This behaviour is most conveniently shown by considering the chemical potentials, activities or activity coefficients of the components.

The chemical potential for ideal solutions is given by IV. (57). For a dilute solution the potential must vary with concentration in the same way. Since $x \ll 1$, $d \ln(1 - x) = -dx$, and we can write

$$d\mu_1 = RT d \ln x, \quad (1)$$

$$d\mu_0 = -RT dx. \quad (2)$$

These equations are limiting equations which apply to constant temperature and pressure with an accuracy increasing with increasing dilution. In order to emphasise their limits of validity they can be written in the form

$$\lim_{x \rightarrow 0} \left(\frac{\partial \mu_1}{\partial x} \right)_{T, p} = \frac{RT}{x}, \quad (3)$$

$$\lim_{x \rightarrow 0} \left(\frac{\partial \mu_0}{\partial x} \right)_{T, p} = -RT. \quad (4)$$

Integration of (1) and (2) gives

$$\mu_1 - \mu_{1(1)} = RT \ln x + k_1, \quad (5)$$

$$\mu_0 - \mu_{0(0)} = -RTx, \quad (6)$$

The presence of the undetermined integration constant in (5) is due to the fact that the concentration $x = 1$ corresponding to the potential $\mu_{1(1)}$ is outside the dilute range in which (1) and (2) are valid. This does not apply to the concentration $x = 0$, so that there is no integration constant in (6).

The use of the activity function to describe dilute solutions leads to relations of a still simpler form. According to [IV. 3. e] the activity in ideal solutions is proportional to the concentration x . The limiting laws in the dilute range can therefore be written as

$$a_1 = k_1 x, \quad (7)$$

$$a_0 = k_0 (1 - x), \quad (8)$$

where k_1 and k_0 are constants. k_1 can be chosen arbitrarily, while

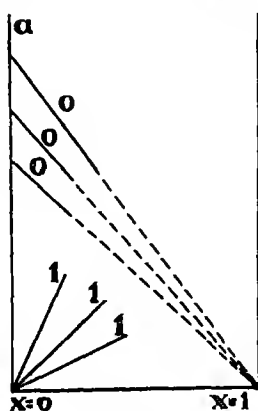


Fig. 1.

k_0 is of course given by $k_0 = a_{0(0)}$. The meaning of these equations is shown in Fig. 1, where the straight lines marked 1 and 0 are tangents to the $a_1 - x$ curve and the $a_0 - x$ curve respectively at $x = 0$. While the direction of the a_1 -curve at $x = 0$ is undetermined, equation (8) shows that the a_0 -curve at $x = 0$ must point in the direction of the point $x = 1$ on the abscissæ axis.

Finally the laws of dilute solution can be expressed in terms of the activity coefficients, defined by equation IV. (63),

$$a_1 = x f_1, \quad (9)$$

$$a_0 = (1 - x) f_0. \quad (10)$$

If (7) is the limiting law, we can write in the neighbourhood of $x = 0$

$$a_1 = k_1 x + k_2 x^n,$$

where $n > 1$. This is equivalent to

$$f_1 = k_1 + k_2 x^{n-1},$$

whence

$$\frac{df_1}{dx} = k_2(n-1)x^{n-2},$$

or

$$x \frac{df_1}{dx} = k_2(n-1)x^{n-1}.$$

The limiting law at infinite dilution is thus

$$x \frac{df_1}{dx} = 0. \quad (11)$$

Similar operations with a_0 give as the limiting law for $x = 0$,

$$\frac{df_0}{dx} = 0. \quad (12)$$

This shows that f_0 , the activity coefficient of the solvent, is independent of the concentration in the dilute range. Since (12) is only a limiting law, it is more correctly expressed by saying that f_0 is represented by a curve which has a tangent parallel to the x -axis at $x = 0$. Equation (11) shows that the corresponding relation does not hold for f_1 .

Equation (11) can also be written

$$x \frac{d \ln f_1}{dx} = 0, \quad (13)$$

whence

$$\frac{d \ln f_0}{dx} = 0, \quad (14)$$

can be derived by means of IV. (68).

Equations (13) and (14) (or the identical pair (11) and (12)) are thus thermodynamically interrelated by the Gibbs-Duhem equation expressed by IV. (68). The same is of course true of equations (7) and (8), which formulate the laws of dilute solutions in terms of activities. The Gibbs-Duhem equation can also be written in the form

$$x \frac{d \ln f_1}{dx} + (1-x) \frac{d \ln f_0}{dx} = 0, \quad (15)$$

where for solutions in general it is only necessary that the sum of the two terms should be zero. In dilute or ideal solutions, however, each of the terms in equation (15) must separately be zero.

For dilute solutions of non-electrolytes f_0 will often be represented by

$$\ln f_0 = \alpha x^2,$$

which according to (15) corresponds to

$$\ln f_1 = -2\alpha x.$$

If the solvent curve is represented by a parabola, the solute curve will thus be a straight line, and *vice versa*. This is illustrated in Fig. 2, where the ordinate is $\ln f$. For normal liquids these expressions are fairly often approximately valid over the whole concentration range. Thus the figure represents the behaviour of solutions of carbon disulphide in methyl alcohol, or of chloroform in acetone. At ordinary temperatures the values of α are 0.54 for the first mixture and -0.43 for the second mixture, corresponding to positive and negative vapour pressure curves respectively. The above formulæ are also applicable to certain alloys, *e.g.*, with $\alpha = 1.0$ they represent the behaviour of the tin-thallium system at 350°.

In order to fix the numerical values of the activity coefficients [IV. 3. e] we have used the usual convention that $f_0 = f_1 = 0$ for $x = 0$, *i.e.*, when the solute

is present at infinite dilution, so that in this range the activity is equal to the concentration for both components. This convention does not however fix the concentration scale for the solute, which can be chosen arbitrarily. The essential content of the laws developed for dilute solutions is not of course affected by this choice.

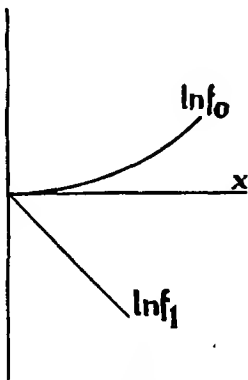


Fig. 2.

2. VAPOUR PRESSURE RELATIONS

(a) **Henry's Law.** We will consider a dilute solution of a volatile substance K_1 in a solvent K_0 . The vapour pressure is always proportional to the activity [IV. (77)], and for dilute solutions the activity is proportional to the concentration (equation (7)) : hence we have

$$p_1 = k_1 x. \quad (16)$$

The partial pressure of the solute at constant temperature is thus directly proportional to its concentration. This is *Henry's law*.

In dilute solutions at constant temperature and pressure the mole fraction scale and the volume concentration scale are directly proportional to one another, so that (16) can also be written as

$$p_1 = k'c. \quad (17)$$

It is easily seen that an analogous argument can be used to derive (7) from (16). This means that the validity of Henry's law is equivalent to the validity of the basis for the laws of dilute solution laid down in the preceding section.

The content of (16) and (17) can also be expressed by saying that the solubility of a gas in a liquid is proportional to the pressure of the gas. The solubility of the gas can be expressed in terms of the *absorption coefficient*, i.e., the amount of substance dissolved in unit volume under unit pressure. In calculating this coefficient the pressure must of course be small enough for (16) and (17) to be valid.

For *Bunsen's absorption coefficient* α the amount of substance dissolved in unit volume under a pressure of one atmosphere is expressed as the volume of gas reduced to 0° and 760 mm. Hg, while *Ostwald's solubility coefficient* l is the ratio between the volume concentrations in the liquid and gas phases. It is easily seen that the two coefficients are related to c , the molar concentration of the solution, by the equations

$$l = \alpha \frac{T}{273.1} = 0.0821 T \frac{c}{p}. \quad (18)$$

Henry's vapour pressure law can be derived from simple kinetic considerations. We shall assume that in a system of two components (just as for a pure substance) the equilibrium between liquid and vapour can be considered as a state in which an equal number of molecules of the volatile substance are vaporised and are condensed in unit time. Both these numbers will in general only represent a fraction of the total number of molecules striking the surface, but provided that the system is so dilute that there is no interaction between K_1 -molecules either in the gas phase or the

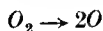
solution, then this fraction will be independent of the concentration. If therefore the concentration in the gas phase is altered in a certain ratio, the concentration in the solution must be altered in the same ratio in order to maintain equilibrium. This simple kinetic treatment thus leads directly to Henry's law, and hence in the light of what has been said above may be considered as the basis of the theory of dilute solutions.

The derivation of (16) and (17) shows that they are only valid when the pressure and concentration refer to the same molecular species. The validity of Henry's law thus depends upon the fact that the solute has the same molecular weight in the gaseous and the dissolved states, so that it is possible in principle to determine molecular weights in solution on the basis of Henry's law. If the solute is partly associated or dissociated in solution so that it may be said to contain both "single molecules" and "double molecules," then the above equations can only be applied to each of these species separately. The relation between the total solute vapour pressure and its concentration depends on the relation between the concentrations of the two species, which in such cases can be calculated by means of the law of mass action [VI. 2. 3].

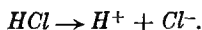
If the solute molecule is completely dissociated into two molecules in solution, the law of mass action shows that the relation between its vapour pressure and its concentration in solution is given by

$$p_1 = kc^2. \quad (19)$$

This equation is valid for solutions of oxygen in molten silver, and approximately for solutions of hydrogen chloride in water, the reactions taking place in the two cases being



and



(b) **Raoult's Law.** This law deals with the variation in the vapour pressure of the solvent K_0 when the concentration x varies at constant temperature. Since a_0 is proportional to p_0 [IV. (77)] and also to $(1 - x)$ (equation (8)), we can write

$$p_0 = k_0(1 - x),$$

or

$$\frac{dp_0}{dx} = -p_0, \quad (20)$$

where p_0 is the vapour pressure of pure K_0 . For finite concentrations we have approximately

$$\frac{\Delta p_0}{p_0} = -x. \quad (21)$$

This equation expresses Raoult's law. If the solvent is water it can be written

$$\frac{\Delta p_0}{p_0} = -\frac{c}{55.5}, \quad (22)$$

where c is the molarity, i.e., the number of gram-molecules per 1,000 grams of solvent.

The law thus states that the *relative vapour pressure lowering* caused by dissolving a small quantity of substance in a solvent is equal to the concentration of the solution formed expressed as the mole fraction x .

Since the molecular weights of K_0 and K_1 are involved in x , equation (21) can be used for determining molecular weights. If the solution is composed of g_0 and g_1 grams of the two components, and the molecular weights are M_0 and M_1 respectively, we have for sufficiently dilute solutions,

$$x = \frac{n_1}{n_0} = \frac{g_1 M_0}{g_0 M_1}, \quad (23)$$

whence the ratio M_0/M_1 can be calculated if x is determined from vapour pressure measurements according to (21).

Since the molecular weights are not necessarily the same in the two states of aggregation, it is necessary to investigate further the significance of the method for determining molecular weights provided by (23). For this purpose we shall return to equation (2) on which Raoult's law is based, i.e.,

$$d\mu_0 = -RTdx.$$

If c is the molar concentration of K_1 and $V_{0(0)}$ the molar volume of K_0 , we have

$$x = V_{0(0)} c \quad (24)$$

which in combination with (2) gives

$$d\mu_0 = -RTV_{0(0)} dc. \quad (25)$$

Since the ratio of μ_0 to $V_{0(0)}$ is independent of the molecular weight assigned to K_0 , the last equation shows that no determination of M_0 is possible on the basis of (2). On the other hand, since μ_0 and $V_{0(0)}$ refer to the same arbitrary amount of solvent, (25) can be used to determine dc , the volume concentration of the solute, from which M_1 can be calculated if the weight composition of the solution is known.

Raoult's law is obtained directly by introducing the equation valid for one mole of a gas, II. (31), i.e.,

$$d\mu_0 = RTd\ln p_0.$$

Raoult's law is thus based on the molecular weight of K_0 in the gas state, and the quantity M_0 occurring in (23) refers to this state.

Thus while Henry's law gives the molecular weight of the solute in terms of its molecular weight in the gas state, M_1 , the application of Raoult's law gives it in terms of M_0 , the molecular weight of the solvent in the gas state. M_1 can also be obtained directly from (2) or (25) without any such assumptions. It should however be emphasised that what is obtained in the first instance is always the *number of molecules*, from which the *molecular weight* is calculated on the basis of a known amount of substance. These methods therefore give no information as to how far the solute molecules are present as such in solution, or whether they are combined with solvent molecules. The combination of solutes with the solvent is known as *solvation*, or as *hydration* in the case of aqueous solutions. It probably occurs in many cases, but is difficult to detect or measure with certainty.

Equations (21) and IV. (81) show that solutions obeying Raoult's law will exhibit no evolution or absorption of heat when diluted with the solvent.

(c) **Elevation of Boiling Point.** Henry's law and Raoult's law for the partial vapour pressures of dilute solutions are valid independent of whether only one or both components are volatile. On the other hand, the simple law derived below for the boiling point will only be valid if the solvent (and not the solute) can pass into the vapour phase.

Fig. 3 shows the vapour pressure curve for the pure solvent K_0 ,

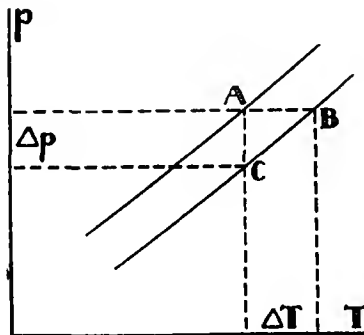


Fig. 3.

passing through A , and the vapour pressure curve for the solution consisting of K_0 plus a small amount of non-volatile K_1 . Since the vapour pressure of the solution is lower than that of the pure

solvent, its boiling point must be higher. AC represents the vapour pressure lowering, and AB the boiling point elevation. It will be seen from the figure that if the two curves are close together (*i.e.*, if the solution is dilute) the ratio of the vapour pressure lowering to the boiling point depression is given by

$$\frac{AC}{AB} = - \frac{\Delta p}{\Delta T} = \frac{dp}{dT},$$

whence

$$\Delta T = - \frac{dT}{dp} \Delta p. \quad (26)$$

The boiling point elevation is thus determined by the vapour pressure lowering combined with the slope of the vapour pressure curve of the solvent.

It is convenient to convert (26) to a somewhat different form by introducing II. (48). We then obtain

$$\Delta T = - \frac{Tv}{q} \Delta p, \quad (27)$$

whence by introducing Raoult's law, (21),

$$\Delta T = \frac{Tvp_x}{q}$$

or

$$\Delta T = \frac{RT^2 x}{q}. \quad (28)$$

This equation can also be derived by using the expressions for the variation of the chemical potential with temperature and with concentration, analogously to the method used for obtaining equation (40).

According to (28) the boiling point elevation is determined by the boiling point of the solvent, T , its molar heat of vaporisation, q , and the concentration of the solute. On the other hand, the nature of the solute does not enter into the equation.

If the solvent is water, (28) can be re-written as

$$\Delta T = \frac{RT^2 c}{q \cdot 55.5}, \quad (29)$$

or, putting $T = 373$ and $q = 536 \times 18$,

$$\Delta T = 0.52c. \quad (30)$$

The factor multiplying c is known as the *molar boiling point elevation*.

It cannot actually be measured in a 1-molar solution, since such a solution is not sufficiently dilute for the equation to be valid.

Equation (28) can be used for determining the molecular weights of substances in solution, just as in the case of vapour pressure lowerings. Since the boiling point of a solution can be determined fairly easily experimentally, the method is much used for molecular weight estimations.

(d) **The Distribution Law.** If a third substance is added to a system consisting of two incompletely miscible liquids in equilibrium, it will in general distribute itself between the two liquid layers. If the added substance is present in small amount so that the laws of dilute solutions can be applied, then the distribution ratio will be constant and independent of the actual quantities present.

This is easily shown by applying Henry's law to the equilibrium between the two liquid phases *A* and *B* and the common vapour phase. The partial vapour pressure of the added substance is then given by (17) as

$$p = k_A c_A,$$

$$p = k_B c_B,$$

whence we obtain the distribution law,

$$\frac{c_A}{c_B} = \frac{k_B}{k_A} = \text{constant.} \quad (31)$$

It can of course also be obtained by direct use of (1) for the potentials.

3. FREEZING POINT AND SOLUBILITY RELATIONS

(a) **The Solubility of Solids.** Equations IV. (87) and IV. (92) give general expressions for the effect of pressure and temperature on the solubility of solids. These equations cannot in general be used for any quantitative calculations unless the dependence of the potential on the concentration is known. However, if the solid in question is sparingly soluble, so that the saturated solution is dilute, then this dependence is given by equation (3) and the calculations can be carried out.

If (3) is introduced into IV. (87) we obtain

$$\left(\frac{\partial x}{\partial p}\right)_T = \frac{V'_1 - V_1}{RT}, \quad (32)$$

or

$$RT \frac{d \ln x}{dp} = V'_1 - V_1 = -\Delta V_1, \quad (33)$$

where ΔV_1 is the increase of volume accompanying the process of solution. This is in general so small that in practice very large pressures are necessary to produce any measurable changes in solubility.

Introducing (3) in IV. (92), we obtain

$$\left(\frac{\partial x}{\partial T}\right)_p = \frac{S_1 - S'_1}{\frac{RT}{x}}, \quad (34)$$

or

$$RT \frac{d \ln x}{dT} = S_1 - S'_1, \quad (35)$$

or by introducing IV. (93),

$$RT^2 \frac{d \ln x}{dT} = q. \quad (36)$$

The method of derivation shows that q is the differential heat of solution for one gram-molecule of K_1 in the saturated solution. For dilute solutions this heat of solution is identical with the irreversible heat of solution for the pure solvent. If the heat of solution per gram is known from calorimetric experiments, then since q is the molar heat of solution the molecular weight can be determined from equation (36).

(36) has the same form as the equation for the vapour pressure of a pure liquid, and if q , the heat of solution, is negative, the solubility curve will have a form similar to that of a vapour pressure curve. If the solubility is very small, the molar heat of solution will as a rule have a large negative value.

For sparingly soluble substances it is possible to calculate the ratio of the solubilities of different states of aggregation or different allotropic modifications. Thus if s_1 and s_2 are the solubilities of two allotropic modifications of the same substance, we can replace the vapour pressures in II. (52) by solubilities, giving

$$A = RT \ln \frac{s_1}{s_2}. \quad (37)$$

It is thus possible to calculate the solubility ratio from the work of transfer, A , or *vice versa*. The more precise equation

$$\mu_1 - \mu_2 = RT \ln \frac{a_1}{a_2}, \quad (38)$$

where a_1 and a_2 are the activities of the two modifications is obtained directly from IV. (61).

(b) **Freezing Point Depression.** In Fig. 4 V and K are the vapour pressure curves for the solvent K_0 in the liquid and solid states respectively, while the curve O is the vapour pressure curve of the liquid solution. The points of intersection P and P' at which the vapour pressures of the liquid and solid are equal are the freezing

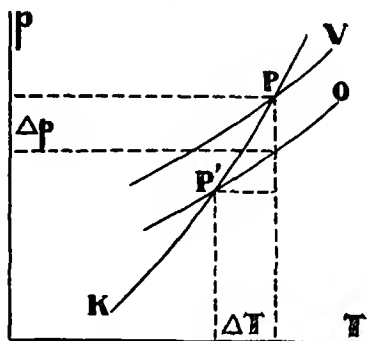


Fig. 4.

points of pure solvent and solution respectively. It is seen from the figure that the solution has a lower freezing point than the solvent, the depression of freezing point $-\Delta T$ corresponding to the vapour pressure lowering $-\Delta p$.

It is also clear from the figure that if the solution is sufficiently dilute we can write

$$\frac{\Delta p}{\Delta T} = \frac{dp_2}{dT} - \frac{dp_1}{dT},$$

where p_2 and p_1 are the vapour pressures of the solid and the solution respectively. Introducing II. (53)

$$q = T\nu \left(\frac{dp_2}{dT} - \frac{dp_1}{dT} \right),$$

and Raoult's law (21), we obtain

$$-\Delta T = \frac{RT^2x}{q}, \quad (39)$$

where q is the molar heat of fusion of K_0 . For a given solvent ΔT thus depends only on x and not on the nature of the solute.

This expression for the freezing point depression can also be obtained by applying IV. (92) to K_0

$$\left(\frac{\partial x}{\partial T}\right)_p = \frac{S_0 - S'_0}{\left(\frac{\partial \mu_0}{\partial x}\right)_{T,p}}$$

and introducing (4),

$$\left(\frac{\partial \mu_0}{\partial x}\right)_{T,p} = -RT.$$

This gives

$$-\left(\frac{\partial x}{\partial T}\right)_p = \frac{S_0 - S'_0}{RT},$$

or by introducing IV. (93)

$$-\frac{x}{\Delta T} = \frac{q}{RT^2},$$

identical with (39).

If the molarity concentration scale is introduced in (39), we have for aqueous solutions

$$-\Delta T = \frac{RT^2 c}{q \cdot 55.5}, \quad (40)$$

where the factor multiplying c is termed the *molar freezing point depression*. For water we have $T = 273.1$ and $q = 18 \times 80$, giving

$$-\Delta T = 1.86c. \quad (41)$$

Equations (28) and (39) show that the addition of another substance affects the boiling point and the freezing point in opposite directions, raising the former and lowering the latter. This is because the phase in which the added substance dissolves is transformed into the coexistent phase with absorption of heat in the first case and evolution of heat in the second case.

Like the elevation of boiling point, the depression of freezing point can be used for determining molecular weights, and it is particularly suitable for this purpose since the estimation can be carried out with great accuracy.

The assumption underlying the above derivations is that the freezing point of the solution is determined by its equilibrium with the solvent in the pure crystalline state: in other words, it is assumed that the added substance K_1 dissolves only in the liquid. If it is also dissolved by the solid phase the freezing point will be determined by the intersection of the vapour pressure curves of the liquid and solid solutions, O_1 and O_2 respectively in Fig. 5. Both these curves lie below the curves V and K for the vapour pressures of pure solvent in the liquid and solid states respectively.

It is easily seen from the figure that the freezing point of the solution, T_2 , need not be lower than that of the pure solvent, but that their relative positions depend on the relative vapour pressure lowerings in the two states of aggregation. If $-\Delta p$ is greater for the liquid,

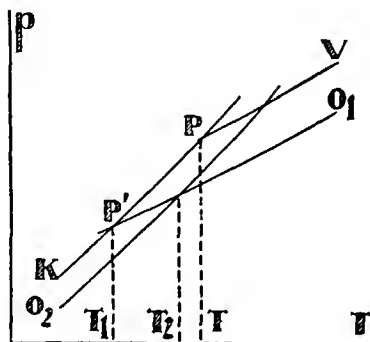


Fig. 5.

the freezing point is lowered, while if $-\Delta p$ is greater for the solid the freezing point is raised: finally, if Δp is the same in the two phases, the freezing point will not be affected by the addition of solute.

If Raoult's law is assumed to hold both for the liquid and the solid phases, then the freezing point will be lowered or raised according as K_1 is more soluble in liquid or solid K_0 . This statement agrees with the law given in [IV. 4. b] for the melting points of mixed crystals.

4. OSMOTIC PRESSURE

(a) **Definition and Measurement.** If a solution and pure solvent are at the same pressure and temperature they cannot be in equilibrium with one another, since we shall always have

$$\mu_1 > \mu_{1(0)},$$

$$\mu_0 < \mu_{0(0)}.$$

It is however possible to establish partial equilibrium by increasing the pressure on the solution. This will cause μ_0 to increase according to equation IV. (52), which can be written

$$\left(\frac{\partial \mu_0}{\partial p}\right)_{T,x} = V_0, \quad (42)$$

where V_0 is the differential molar volume of the solvent. If the

solution is dilute, V_0 can be identified with $V_{0(0)}$, the molar volume of the pure solvent.

The excess pressure which must be applied to the solution in order that the condition

$$\mu_0 = \mu_{0(0)}, \quad (43)$$

shall be fulfilled is termed the *osmotic pressure* of the solution. It can be defined as the difference of pressure which must be set up between the solution and the solvent in order that the system shall be in equilibrium with respect to K_0 . If K_1 is prevented from passing from the solution to the solvent, this represents a state of complete equilibrium.

There are various ways of observing when equilibrium has been set up; e.g., by means of a semi-permeable membrane which allows the passage of the solvent K_0 but not of the solute K_1 . If there is no movement of liquid through the membrane, and small changes of pressure cause movement in either direction, then the system is said to be in osmotic equilibrium and the osmotic pressure can be measured directly by means of a manometer.

The following arrangement can be used for carrying out measurements in practice. A semi-permeable cylinder with a vertical tube attached is filled with the solution and immersed in pure solvent K_0 . The tendency for the solution to become diluted will force K_0 into the cylinder, thus causing the liquid to rise in the vertical tube until the osmotic pressure has been attained. The pressure is measured by the height of the liquid in the tube and its density. (Cf. Fig. 6.)

For aqueous solutions the semi-permeable membrane can consist of cupric ferrocyanide deposited in the pores of a clay cylinder: this may be used, for example, to measure the osmotic pressure of a solution of cane sugar. Membranes of gelatine or of animal origin can also be used.

Measurements of osmotic pressure in dilute solutions have shown that its dependence on the concentration and the temperature is expressed by the equation

$$P = RTc. \quad (44)$$

Since the concentration c is the reciprocal of V' , the volume of solution containing one gram-molecule of K_1 , this equation can also be written

$$PV' = RT, \quad (45)$$

showing that the osmotic pressure follows the same laws as the pressure of a gas. This law was enunciated by van't Hoff, and was the original basis on which the laws of dilute solution were developed.

If the solute is non-volatile, then instead of using a semi-permeable membrane to establish osmotic equilibrium, we can use the passage of the solvent into the vapour state. Further, a positive pressure on the solution can be replaced by a negative pressure (or tension) on the solvent. The vaporisation of the solvent under a negative pressure can be realised by using a capillary tube or a porous plate with capillary pores.

Equations (44) and (45) can be deduced directly on the basis of the theory of dilute solutions, *e.g.*, from equation (4). If there is equilibrium at constant temperature between solvent and solution, then the decrease in the chemical potential of the solvent brought about by the addition of K_1 must be just compensated by the increase caused by the increased pressure. We thus have as the condition for equilibrium

$$d\mu_0 = \left(\frac{\partial\mu_0}{\partial x}\right)_{T,p} dx + \left(\frac{\partial\mu_0}{\partial p}\right)_{T,x} dp = 0,$$

whence by introducing (4) and (42)

$$RTdx = V_0 dp,$$

which on combining with (24) gives

$$dp = RTdc,$$

identical with (44).

The osmotic pressure can be used for determining the molecular weight of the solute, since equation (44) gives the number of gram-molecules of K_1 present in unit volume. The direct applicability of osmotic pressures for this purpose depends on the significance of (25) as a basis for the laws of osmotic equilibrium.

(b) **Osmotic Work.** If a semi-permeable piston is moved in opposition to the osmotic pressure, it is possible to isolate pure solvent from a solution. If we assume that in order to isolate one gram-molecule of K_0 it is necessary to move the piston through a volume V'' , the volume of solution containing one gram-molecule of K_0 , then the work needed is

$$A_0 = PV'', \quad (46)$$

which must be identical with the differential work of mixing. Since the volumes containing one gram-molecule of the two components are in the inverse ratio of their concentrations, we have

$$\frac{V'}{V''} = \frac{1-x}{x}, \quad (47)$$

and hence, since $x \ll 1$,

$$A_0 = PV'x, \quad (48)$$

or, introducing (45),

$$A_0 = RTx. \quad (49)$$

In this derivation we have neglected the compressibility of the solution and the change in total volume when K_0 is separated from the mixture: this is permissible for sufficiently dilute solutions.

If (49) is inserted in the general equation IV. (50)

$$A_0 = -(\mu_0 - \mu_{0(0)}) + p(V_0 - V_{0(0)}),$$

at the same time (since $x \ll 1$) putting

$$V_0 = V_{0(0)},$$

we obtain

$$\mu_0 - \mu_{0(0)} = -RTx, \quad (50)$$

which is identical with the equation (6) which was previously regarded as the basis for the equations derived for dilute solutions. The theory of dilute solutions can therefore be based equally well on (49), derived by means of the osmotic pressure, or on (6), derived from the potential function. It must however be observed that both these equations represent limiting laws which are in general exactly valid only at infinite dilution.

(c) **Osmotic Pressure and Vapour Pressure.** It has been previously pointed out that the laws of dilute solution in their simple form (e.g., Raoult's law (21)) are not valid if the forces between the molecules are of such a kind that association or dissociation reactions take place. On the other hand, purely thermodynamic relations will be unaffected by such complications. For example, we can derive a relation between the osmotic pressure and the vapour pressure of a dilute solution as follows, without assuming the validity of the gas laws.

Putting $V'' = V_{0(0)}$ in (46), we have

$$d\mu_0 = -dA_0 = -V_{0(0)}dP, \quad (51)$$

or

$$\mu_0 - \mu_{0(0)} = -V_{0(0)}P.$$

Further, from II. (31)

$$d\mu_0 = RTd\ln p_0 = v_0 dp_0,$$

or

$$\mu_0 - \mu_{0(0)} = v_0 \Delta p_0,$$

where $-\Delta p_0$ is the vapour pressure lowering. We thus obtain

$$P = -\frac{v_0}{V_{0(0)}} \Delta p_0. \quad (52)$$

The ratio of the osmotic pressure to the vapour pressure lowering is thus equal to the ratio of the volumes of the solvent in the vapour and liquid states.

The same formula can be derived by considering the osmotic apparatus in Fig. 6. At the surface of the solution in the vertical tube the vapour pressure is $p + \Delta p$, where $-\Delta p$ is the vapour

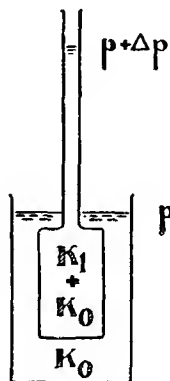


Fig. 6.

pressure lowering, while the vapour pressure at the surface of the solvent is p . A column of liquid of the same height as the column of liquid in the vertical tube is thus supported by the pressure difference $-\Delta p$, while the column of liquid is supported by the osmotic pressures. These two pressures must therefore be in the same ratio as the corresponding densities, *i.e.*,

$$-\frac{\Delta p}{p} = \frac{d_D}{d_V}, \quad (53)$$

where d_D and d_V are the densities of the vapour and the liquid respectively. But for dilute solutions we have

$$\frac{d_D}{d_V} = \frac{V_{0(0)}}{v_0},$$

which on inserting in (53) gives an expression identical with (52).

In a similar way the osmotic pressure can be expressed in terms of the boiling point elevation and the freezing point depression.

5. TERNARY MIXTURES

A ternary mixture in which two of the components predominate can be considered as a dilute solution in a mixed solvent. According

to IV. (68) we have for a mixture of three components at constant pressure and temperature,

$$n_1 d\ln f_1 + n_2 d\ln f_2 + n_3 d\ln f_3 = 0$$

or, if the two components constituting the solvent are termed K_{01} and K_{02} and the solute K_1 ,

$$n_{01} d\ln f_{01} + n_{02} d\ln f_{02} + n_1 d\ln f_1 = 0, \quad (54)$$

or

$$n_0 d\ln f_{01} \frac{n_{01}}{n_0} \frac{n_{02}}{f_{02}} + n_1 d\ln f_1 = 0, \quad (55)$$

where $n_0 = n_{01} + n_{02}$. We thus obtain an equation analogous to the equation for binary mixtures by introducing

$$f_0 = f_{01} \frac{n_{01}}{n_0} \frac{n_{02}}{f_{02}}. \quad (56)$$

Since the ratio of the amounts of the solvent components is kept constant, we can assume that Henry's law will hold in the form

$$x \frac{d\ln f_1}{dx} = 0, \quad (57)$$

where $x = \frac{n_1}{n_1 + n_0}$. If this is so, then Raoult's law

$$\frac{d\ln f_0}{dx} = 0 \quad (58)$$

will also hold for the ternary system if f_0 is defined as above. It is not however possible to derive an analogous equation for the single components of the solvent on the basis of the Gibbs-Duhem equation. It is in fact possible that the addition of K_1 to the mixed solvent may bring about changes of opposite sign in the activity coefficients, activities or potentials of the two other components. Thus if sodium chloride is added to a mixture of alcohol and water the vapour pressure of the alcohol will increase, while that of the water will decrease.

If (57) is assumed to hold, we can write for the variations of f_{01} and f_{02} with x ,

$$\frac{d\ln f_{01}}{dx} = \alpha_1, \quad \frac{d\ln f_{02}}{dx} = \alpha_2, \quad (59)$$

so that the directions of the two curves at $x = 0$ are connected by the equation

$$\frac{\alpha_1}{\alpha_2} = -\frac{n_{02}}{n_{01}}. \quad (60)$$

The above equations will be used later [XI. 2. h] for calculating the osmotic pressure in ternary mixtures.

CHAPTER VI

HOMOGENEOUS EQUILIBRIA

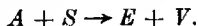
1. BALANCED REACTIONS

IN the foregoing chapters, when we have considered a homogeneous phase (either alone or as part of a heterogeneous system), we have assumed that its properties are completely determined by its empirical composition and the external conditions of temperature and pressure.

This is not however always the case. Thus if alcohol and acetic acid are mixed we obtain a homogeneous liquid having properties which alter with the time. The properties of this homogeneous phase under constant temperature and pressure are thus not determined by its empirical composition alone. The changes which its properties undergo must therefore be attributed to spontaneous internal changes, *i.e.*, to chemical processes taking place between the substances of which the phase is constituted. The nature of these changes can often be demonstrated by the isolation of fresh substances from the mixture.

The laws for processes of this kind are entirely different from the laws governing purely heterogeneous processes such as fusion, polymorphic transformations, or the more complex heterogeneous processes dealt with in [IV. 5. a.]. The difference depends upon the fact that when a homogeneous reaction takes place under the conditions stated above, there is a continuous change in the chemical potentials of the reacting substances, while in the heterogeneous reactions the chemical potentials remain constant in spite of changes in the amounts of substance present.

The reaction between alcohol and acid mentioned above leads to the formation of the new substances ester and water, and may be represented by the equation



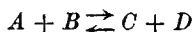
As the reaction proceeds from left to right, the concentrations (and hence the chemical potentials) of the substances *A* and *S* of the initial system will continually decrease, while the concentrations and potentials of *E* and *V*, the substances formed, will continually increase. Both these changes will decrease the tendency of the process to take place from left to right, and as the process proceeds

this tendency must become zero before complete transformation has taken place. If we start with $E + V$, this system will undergo a change obeying analogous laws and resulting in the production of $A + S$: this transformation will also be incomplete. If we start with equivalent quantities of $A + S$ on the one hand and $E + V$ on the other, the final state in which *homogeneous equilibrium* has been reached will be the same in the two cases. An incomplete process of this kind, leading to a state of equilibrium, is termed a *balanced reaction*, and for the case in question is written in the form



It was at one time supposed that balanced reactions represented exceptional cases, but experience has shown that this is not so. The above considerations make it natural to regard all homogeneous processes as balanced reactions, so that those processes usually termed complete must be considered as incomplete processes in which the equilibrium state is for practical purposes identical with one of the systems, *i.e.*, the first product in a pure state.

The idea of incomplete homogeneous processes was first developed by Berthollet in 1801. The theory was developed further and formulated quantitatively by Guldberg and Waage in 1867. According to the latter authors the "action" of a substance in a homogeneous mixture is proportional to its "active mass" or concentration. If we have an equilibrium between four substances expressed by



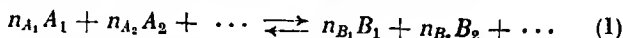
an increase in the concentration of A or of B will displace the equilibrium in favour of the system $C + D$, while an increase in the concentrations of the last two substances will cause a displacement in the direction of $A + B$. Since the "action" is proportional to the concentration c , the tendency for the process from left to right will be proportional to the product $c_A c_B$, while that of the reverse process is proportional to $c_A c_D$. Once equilibrium has been reached the ratio of these products must not be displaced, and we therefore have as the equilibrium condition

$$\frac{c_C c_D}{c_A c_B} = \text{constant}.$$

This is the mathematical expression for the *law of mass action* applied to a reaction taking place according to the above scheme. The above derivation is however unsatisfactory on account of the vague nature of the undefined expressions "action" and "tendency."

2. THE LAW OF MASS ACTION

(a) **Thermodynamic Derivation.** To obtain a thermodynamic deduction of the law of mass action we shall first consider a balanced reaction involving gaseous substances obeying the perfect gas laws. The general scheme for such a process can be written



We now imagine this reaction to take place *completely*, reversibly and isothermally from left to right, the reacting substances being taken from or added to infinitely large reservoirs in which the concentrations are C_{A_1} , C_{A_2} , . . . and C_{B_1} , C_{B_2} , We also imagine the same substances contained in a reaction vessel of constant volume in which chemical equilibrium has been set up, and in which the concentrations are c_{A_1} , c_{A_2} , . . . and c_{B_1} , c_{B_2} , If the amounts of the initial substances represented by the left-hand side of equation (1) are transferred reversibly from the reservoirs to the reaction vessel, the work obtained is

$$n_{A_1}RT \ln \frac{C_{A_1}}{c_{A_1}} + n_{B_2}RT \ln \frac{C_{A_2}}{c_{A_2}} + \dots$$

When the A -system is converted to the B -system in the reaction vessel no work can be obtained, since the conversion takes place in an equilibrium state at constant volume. If the resultant substances formed are removed and transferred to the appropriate reservoirs, the work obtained is

$$n_{B_1}RT \ln \frac{c_{B_1}}{C_{B_1}} + n_{B_2}RT \ln \frac{c_{B_2}}{C_{B_2}} + \dots,$$

so that the total work obtained in the whole transformation is

$$A = RT \left[\ln \frac{C_{A_1}^{n_{A_1}} \cdot C_{A_2}^{n_{A_2}} \dots}{C_{B_1}^{n_{B_1}} \cdot C_{B_2}^{n_{B_2}} \dots} + \ln \frac{c_{B_1}^{n_{B_1}} \cdot c_{B_2}^{n_{B_2}} \dots}{c_{A_1}^{n_{A_1}} \cdot c_{A_2}^{n_{A_2}} \dots} \right]. \quad (2)$$

According to the second law, A depends only on the initial and final states of the process. The first term in (2) is a constant containing only the initial and final concentrations; hence the second term containing the variable equilibrium concentrations must also be a constant. We thus have

$$\frac{c_{B_1}^{n_{B_1}} \cdot c_{B_2}^{n_{B_2}} \dots}{c_{A_1}^{n_{A_1}} \cdot c_{A_2}^{n_{A_2}} \dots} = K_c. \quad (3)$$

where K_c is a constant known as the *equilibrium constant*. This is the general expression for the law of mass action; it is also known as the *reaction isotherm*, as it gives the relation between the concentrations at constant temperature.

If we put $C = 1$, then

$$A = RT \ln K_c. \quad (4)$$

The method of derivation shows that the above expression for the law of mass action depends upon the validity of the gas laws for the substances taking part in the reaction. Without making any further assumptions we can therefore introduce in (3) the expression II. (6),

$$p = RTc,$$

giving

$$\frac{P_{B_1}^{n_{B_1}} \cdot P_{B_2}^{n_{B_2}} \cdots}{P_{A_1}^{n_{A_1}} \cdot P_{A_2}^{n_{A_2}} \cdots} = K_p, \quad (5)$$

where

$$K_p = K_c (RT)^{\sum n_{B_i} - \sum n_{A_i}} \quad (6)$$

Instead of using the transference process described above, the law of mass action can be derived in a simple manner by means of thermodynamic functions. According to I. (25), the general condition for chemical equilibrium expressed in terms of the thermodynamic potential is

$$(dG)_{T,p} = 0.$$

When the process (1) takes place in the homogeneous mixture, the variations in the amounts of substances present can be expressed by

$$dn_{A_i} = -n_{A_i} d\alpha \cdots, \quad dn_{B_i} = n_{B_i} d\alpha \cdots. \quad (7)$$

If the initial system is pure A , then α will be the extent of the reaction, i.e., a factor which varies from 0 to 1 as the A -system is transformed to the B -system, and which is proportional to the amount of the latter system. From equations (7) and I. (44), dG in the mixture can therefore be written as

$$dG = -\sum n_{A_i} \mu_{A_i} d\alpha + \sum n_{B_i} \mu_{B_i} d\alpha,$$

giving as the equilibrium condition

$$\left(\frac{dG}{d\alpha} \right)_{T,p} = -\sum n_{A_i} \mu_{A_i} + \sum n_{B_i} \mu_{B_i} = 0. \quad (8)$$

It follows from this equation that the standard states of the

substances taking part in the equilibrium cannot be chosen arbitrarily; thus, for example, if the numerical value of one of the potentials is altered by altering the appropriate standard state, this will not in general be consistent with (8). In other words, the standard states must be chosen in such a way that (8) is satisfied for a *single* equilibrium state: if this is done, then the equation will be satisfied for all equilibrium states at the same pressure and temperature.

It is inherent in the problem that this limitation in the choice of standard states is present for all reactions which involve stoichiometric relations between the components, and it has previously been introduced as a self-evident condition, *e.g.*, in the distribution of a substance between two phases, and other simple cases. The same result emerges from the phase rule by considering the definition of a component given in [IV. 6].

Since the form of a relation between observable quantities cannot be affected by the choice of standard states for the substances concerned, it is important to note that freedom in this choice can be introduced by adding to equation (8) a constant term, the value of which is fixed together with the selected standard states on the basis of a single equilibrium state of the system. On this basis the condition for equilibrium becomes

$$-\sum n_{A_i} \mu_{A_i} + \sum n_{B_i} \mu_{B_i} = I, \quad (9)$$

where I is a constant independent of temperature and pressure. There will of course also be a corresponding alteration in equation I. (35).

We can now introduce IV. (13), which can be written

$$\mu_{A_i} = RT \ln p_{A_i} + \mu_{A_i(1)},$$

where $\mu_{A_i(1)}$ is the potential of A_i under unit pressure. Combining this with (9), we obtain

$$RT[-\sum n_{A_i} \ln p_{A_i} + \sum n_{B_i} \ln p_{B_i}] = I + i, \quad (10)$$

where $i = \sum n_{A_i} \mu_{A_i(1)} - \sum n_{B_i} \mu_{B_i(1)}$ depends only on the temperature. This equation is clearly identical with the law of mass action (5), since $RT \ln K_p = I + i$.

The effect of a change of volume on the equilibrium state can be determined by expressing the concentrations in (3) as ratios between the number of molecules and the volume, *i.e.*,

$$c = \frac{N}{V}.$$

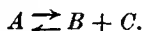
Introducing this relation into (3), we obtain

$$\frac{N_{B_1}^{n_{B_1}} \cdot N_{B_2}^{n_{B_2}} \dots}{N_{A_1}^{n_{A_1}} \cdot N_{A_2}^{n_{A_2}} \dots} V^{\Sigma n_{A_i} - \Sigma n_{B_i}} = K_c, \quad (11)$$

showing that an increase of volume will favour the system which has the greater number of molecules. This is obviously in agreement with le Chatelier's principle.

(b) **Kinetic Derivation.** From the point of view of molecular theory a homogeneous equilibrium (like a vapour pressure equilibrium) cannot be regarded as a static condition in which nothing takes place. It must rather be assumed that equilibrium is set up as a result of balanced processes, in which the molecules of the first system are converted to molecules of the second system at the same velocity with which the reverse process takes place. The laws governing the equilibrium can therefore be derived from the laws for reaction velocity. The detailed consideration of this problem thus belongs properly to reaction kinetics. Since however the equilibrium laws must be independent of any particular reaction mechanism, they may be derived from very simple kinetic considerations, which we shall now describe.

For the sake of simplicity we shall only consider a simple case of homogeneous equilibrium, namely the balanced dissociation equilibrium represented by the scheme



For this process to take place from right to left it is necessary for molecules of B and C to collide, i.e., to be present simultaneously within a very small element of volume. The probability that a B -molecule is present in a given small element of volume is proportional to the number of B -molecules per unit volume, i.e., to the concentration c_B . An analogous statement holds for c_C . The probability that a B -molecule and a C -molecule are simultaneously present in the given volume element is proportional to the product of the two concentrations, since the two contingencies are mutually independent. Since it must also be assumed that at constant temperature a constant fraction of these "collision complexes" react to form A , it is seen that the number of A -molecules formed per unit time in unit volume (i.e., the velocity with which A is formed) must be proportional to the product $c_B c_C$.

The process in the reverse direction can be considered as a spontaneous dissociation of the A -molecule to give the reaction products $B + C$. The probability of this dissociation at a given

temperature is a constant quantity, and the velocity of dissociation is hence proportional to the concentration c_A .

If the two proportionality factors are k_2 and k_1 respectively, we have when equilibrium has been reached,

$$k_1 c_A = k_2 c_B c_C,$$

or,
$$\frac{c_B c_C}{c_A} = \frac{k_1}{k_2},$$

which is identical with the thermodynamic expression (3) applied to the simple case considered here. No difficulties are involved in extending these considerations to a kinetic derivation of the general expression.

(c) **Gaseous Systems.** The above derivations are in the first instance based on a consideration of reactions in the gas phase. We will now consider the application of the law of mass action to some of the simplest special cases of homogeneous equilibria involving gases.

The simplest form which the general equilibrium scheme (1) can assume is

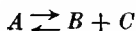


where the equilibrium is established between only two molecules, and may be termed an *isomerisation equilibrium*. In this case the law of mass action (3) reduces to

$$\frac{c_B}{c_A} = K_c, \quad (12)$$

i.e., when equilibrium has been set up, the ratio of the concentrations of the two reacting substances is independent of the total concentration. This equation applies to the equilibria between ortho- and para-hydrogen, and between organic isomers.

Much more important is the reaction between three molecules,



which represents a typical *dissociation equilibrium*. In this case the mass law expression is

$$\frac{c_B c_C}{c_A} = K_c. \quad (13)$$

Suppose we start with pure A at a concentration c_0 , and that when equilibrium has been set up the *degree of dissociation* is α , i.e., α is the fraction of A which is split up into B and C . Then if the dissociation takes place at constant volume, we have

$$c_A = c_0(1 - \alpha), \quad c_B = c_C = c_0\alpha,$$

which on insertion in (13) gives

$$c_0 \frac{\alpha^2}{1 - \alpha} = K_c. \quad (14)$$

Similarly we find for K_p ,

$$p_0 \frac{\alpha^2}{1 - \alpha} = K_p, \quad (15)$$

where p_0 is the pressure of A before dissociation took place. It is obvious that the dissociation causes the total number of molecules to increase in the ratio

$$(1 - \alpha) + 2\alpha = 1 + \alpha$$

and since the pressure at constant volume is proportional to the number of molecules, the pressure p corresponding to α is given by

$$p = p_0(1 + \alpha) \quad (16)$$

which on insertion in (15) gives

$$p \frac{\alpha^2}{1 - \alpha^2} = K_p. \quad (17)$$

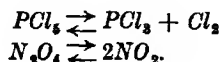
Equation (16) makes it possible to determine α by observing p and calculating p_0 by the equation of state from the amount of substance and the volume. It is clear that (16) can also be expressed in terms of δ and δ_0 , the densities for the dissociated and undissolved states respectively under the same pressure. We have

$$\frac{\delta_0}{\delta} = 1 + \alpha,$$

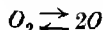
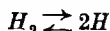
$$\text{or} \quad \alpha = \frac{\delta_0 - \delta}{\delta}. \quad (18)$$

It was this method which first led to the conclusion that substances with abnormally low vapour densities are partly or completely dissociated. It follows from the above formulæ that the addition of one of the dissociation products will repress the dissociation.

As examples of dissociation equilibria for which the law of mass action has been verified, we may mention the dissociation of phosphorus pentachloride and of the double molecules of nitrogen peroxide,



Further examples are provided by the dissociation of the molecules of diatomic elements at high temperatures, e.g.,



An equilibrium between four molecules, e.g., $A + B \rightleftharpoons C + D$, is governed by the mass law expressions

$$\frac{c_C c_D}{c_A c_B} = K_c, \quad (19)$$

or

$$\frac{P_C P_D}{P_A P_B} = K_p. \quad (20)$$

The treatment given above shows that an equilibrium of this kind is independent of the pressure, and further that $K_p = K_c$.

As an example we shall take the reaction $2HI \rightleftharpoons H_2 + I_2$. From a purely chemical point of view this is a dissociation of a compound into its elements, but from the physico-chemical standpoint it cannot be classified with the dissociations dealt with above, since it follows quite different laws: in particular, the degree of dissociation (*i.e.*, the fraction of the original system which has reacted) is independent of the pressure. If the initial system consists of undecomposed hydrogen iodide at a concentration c , and the fraction decomposed at equilibrium is α , then

$$c_{HI} = c(1 - \alpha),$$

$$c_{H_2} = c_{I_2} = c \frac{\alpha}{2},$$

and hence

$$\frac{\alpha^2}{4(1 - \alpha)^2} = K_c. \quad (21)$$

This equation contains no variables: hence the validity of the mass law can only be investigated by varying the relative amounts of iodine and hydrogen, or by adding varying amounts of iodine or hydrogen to dissociating hydrogen iodide.

Starting with a molecules of iodine and b molecules of hydrogen, if at equilibrium γ molecules of each have reacted, then the number of molecules of hydrogen iodide formed is 2γ , and we have

$$\frac{(a - \gamma)(b - \gamma)}{4\gamma^2} = K_c. \quad (22)$$

The following table gives the observed values of 2γ for various

initial amounts of iodine and hydrogen, together with values calculated from the mean value of K_c . The experiments were carried out at the boiling point of sulphur (440°). The agreement between the figures in the last two columns demonstrates the validity of the law of mass action for this reaction.

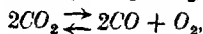
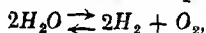
TABLE I. *The Hydrogen Iodide Equilibrium*

a	b	2γ (exp.)	2γ (calc.)
2.94	8.10	5.66	5.64
5.30	7.94	9.52	9.49
9.27	8.07	13.34	13.47
14.44	8.12	14.82	14.93
27.53	8.02	15.40	15.54
33.10	7.89	15.12	15.40

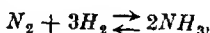
Another reaction which is analogous to the decomposition of hydrogen iodide is the important water-gas reaction.



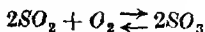
in which the equilibrium is also governed by equation (20). As examples of reactions involving larger numbers of molecules we may mention the decompositions of water vapour and carbon dioxide,



the ammonia equilibrium,



which is of great importance in the manufacture of ammonia, and the formation of sulphur trioxide,



which is also of great industrial importance. All these processes are measurable only at high temperatures and are best studied in the presence of catalysts [VIII. 1. f. and 2. c.].

Some of these equilibria will be treated further in the section dealing with the effect of temperature [4. b.].

(d) **Liquid Systems.** If homogeneous equilibrium has been set up in a gaseous system which is in equilibrium with a liquid, the mass law expression (3) will hold for the gas mixture. Further, if the solution is dilute, Henry's law gives

$$c_{B_1} = k_{B_1} c_{B_1}, \quad c_{B_2} = k_{B_2} c_{B_2}, \dots,$$

where c_B , c_{B_2} , etc., are the concentrations in the gas mixture, and c'_B , c'_{B_2} , etc., the concentrations in the liquid.

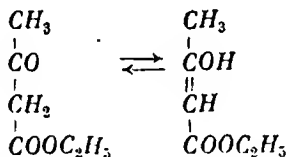
We thus see immediately that equilibria in liquids will be governed by an expression analogous to (3). Since Henry's law was used in arriving at this result, the condition for applying the mass law to a solution is that the gas laws shall be valid.

The law of mass action will therefore be generally valid for dilute solutions. It will however be valid for all concentrations in the case of ideal solutions, where Henry's law holds over the whole concentration range. The law of mass action for liquid systems can also be derived by using the chemical potentials, just as for gases.

In the case of liquids there exist a number of investigations of the simplest type of equilibrium, $A \rightleftharpoons B$, to which equation (12)

$$\frac{c_B}{c_A} = K$$

can be applied. For example, this equation has been found to represent the equilibrium between the two forms of aceto-acetic ester,



and other keto-enol reactions in various solvents. The magnitude of the equilibrium constant depends to a great extent on the medium. Thus in the case of aceto-acetic ester at ordinary temperatures, about 0.5% is enolised in water, and about 46% in hexane, while the pure ester without any solvent is enolised to the extent of about 7%.

An example of a dissociation process is provided by the previously mentioned dissociation of the double molecules of nitrogen peroxide, which also takes place in solution. It can be estimated colorimetrically, since the NO_2 molecule is coloured, while the N_2O_4 molecule is colourless.

The classical example of mass action in liquid systems is the reaction between alcohol and acetic acid, mentioned at the beginning of the chapter, which obeys the mass law approximately. Other examples are provided by the dissociation and other reactions of weak electrolytes, which will be dealt with in a subsequent chapter [VII. 3].

(e) **The Thermodynamic Law of Mass Action.** Equations (3) and (5) depend on the validity of the simple gas laws and are therefore only generally valid for dilute systems, whether gases or liquids. We have seen previously that expressions valid for dilute systems can be transformed into expressions of the same form, but of general validity, by replacing the concentrations by "activities." The expression for the law of mass action can be transformed in the same way.

This can easily be shown by inserting IV. (61)

$$\mu_{A_1} = \mu_{A_1(1)} + RT \ln a_1.$$

in equation (9), giving

$$RT [-\sum n_{A_i} \ln a_{A_i} + \sum n_{B_i} \ln a_{B_i}] = I + i,$$

exactly analogous to (10). This expression can be rewritten in the form

$$\frac{a_{B_1}^{n_{B_1}} a_{B_2}^{n_{B_2}} \dots}{a_{A_1}^{n_{A_1}} a_{A_2}^{n_{A_2}} \dots} = K_a, \quad (23)$$

where K_a is the *thermodynamic mass action constant*. Comparison of (3) and (23) gives the relation between K_a and K_c ,

$$K_c = K_a \frac{f_{A_1}^{n_{A_1}} f_{A_2}^{n_{A_2}} \dots}{f_{B_1}^{n_{B_1}} f_{B_2}^{n_{B_2}} \dots}, \quad (24)$$

where f is the activity coefficient for the temperature and pressure chosen. The derivation of (23) shows that K_a depends on the temperature and pressure, but at constant temperature and pressure is independent of the composition of the mixture. Equation (23) is exactly and generally valid, in contrast to (3) and (5), which depend upon the validity of the gas laws for the components involved. The greater range of validity of (23) depends upon the fact that it is of purely thermodynamic origin, and does not involve the validity of any special equation of state.

3. HOMO-HETEROGENEOUS EQUILIBRIUM

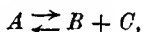
(a) **Thermodynamic Treatment.** A system is said to be in homo-heterogeneous equilibrium when there is chemical equilibrium in one homogeneous phase, and one or more of the components is simultaneously in equilibrium with a second phase. We shall deal particularly with equilibria in which the second phase is a solid.

VI. 3. HOMO-HETEROGENEOUS EQUILIBRIUM 171

The fact that the homogeneous reaction phase is in equilibrium with one of the reactants in the solid state will not affect the validity of the law of mass action for the homogeneous phase. However, the chemical potential and vapour pressure of the solid at constant temperature will be practically constant (the effect of pressure on the potential of a solid being very small), so that equation (10) will still be formally correct if we omit the term referring to this component.

This means that the factors referring to this component can be omitted from the mass action expressions (3) and (5). The mass action expression for a homo-heterogeneous reaction is thus obtained by omitting the factors referring to any substances present as solids. The same result is obtained by using the fact that the vapour pressure or saturation concentration is constant for these components.

If the equilibrium has the simple form



then if either A or B is present as a solid, we have respectively

$$c_A = K_1; \quad c_B c_C = K_2, \quad (25)$$

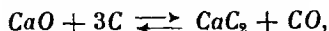
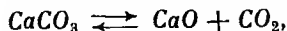
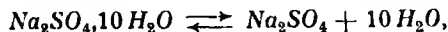
and

$$c_B = K_1, \quad \frac{c_C}{c_A} = K_2, \quad (26)$$

Analogous equations are easily obtained for more complicated cases.

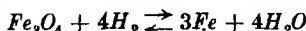
In the case of the thermodynamic law of mass action, the same considerations lead to the conclusion that a homo-heterogeneous equilibrium will be governed by equation (23), with the omission of factors representing the activity of components which are present as pure condensed phases.

(b) **The Equilibrium of Gaseous Systems with Solid Phases.** The simplest case of equilibrium of this type occurs when a single gas is in equilibrium with solid phases. Thus in the following reactions



the mass law predicts that at constant temperature there will be a constant equilibrium pressure of water vapour, carbon dioxide and carbon monoxide respectively. As previously shown, the same result can be obtained for these simple processes by applying the phase rule.

When ferrous-ferrie oxide (Fe_3O_4) is reduced by hydrogen, the reaction scheme is

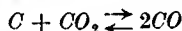


whence by applying the above rules we obtain the mass action expression

$$\frac{c_{H_2O}}{c_{H_2}} = K_c,$$

At 900° C. the value of K_c is about 0.7. If the ratio of the concentrations of the two gases is maintained at a higher or a lower value, complete oxidation of the iron or complete reduction of the oxide will take place.

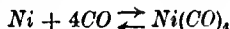
The reaction between carbon dioxide and carbon



is an example of a process obeying the simple dissociation equation

$$\frac{c_{CO}^2}{c_{CO_2}} = K_c.$$

Finally we may mention the formation of nickel carbonyl according to the equation

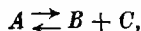


for which the equilibrium expression is

$$\frac{c_{Ni(CO)_4}}{c_{CO}^4} = K_c.$$

On account of the high power to which the carbon monoxide enters in this expression, this equilibrium is particularly sensitive to pressure changes.

(c) **The Equilibrium of Liquid Systems with Solid Phases.** If a solid phase is in equilibrium with a saturated solution in which there is a dissociation equilibrium, *e.g.*,



then as before there are two possible types of equilibrium, according to whether the solid phase is the compound A or one of the dissociation products. If the solubility is low enough for the gas laws to be valid, we shall expect (25) or (26) to hold respectively. (For more concentrated solutions the activities must be inserted in place of the concentrations: *cf.* [2. e.]) In order to investigate this it is necessary to determine the concentration of each molecular species in the solution.

If A is practically completely dissociated in solution, c_B and c_C

will be practically equal to the total concentrations of the substances *B* and *C* as determined by analysis. The curve representing the solubility of *A* will be a rectangular hyperbola if it is plotted with the total concentrations of *B* and *C* as co-ordinates. In the same system of co-ordinates the solubility curves for *B* and *C* will be straight lines. If the dissociation is incomplete, the solution saturated with *A* will also contain *A*-molecules. The usual methods of solubility estimations only give the total amounts of *B* and *C* in the solution, but it is possible to test the law of mass action in this case by comparing the observed solubilities with those calculated theoretically.

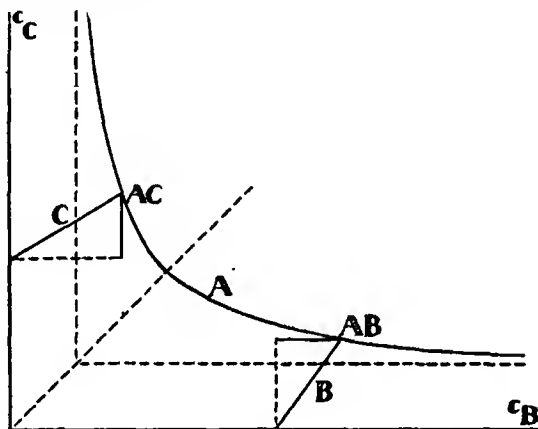


FIG. 1.

In a system of the above type the following solubility phenomena will in general be observed as the amounts of the components are varied. If increasing amounts of *C* are added to the solvent, the solubility of *B* (i.e., the total concentration of *B* in the saturated solution) will increase, since *A*-molecules are formed in solution. At a certain *C*-content the solution will be saturated with *A*, and further addition of *C* will transform all the solid *B* to solid *A*. While this is taking place the solution is saturated with respect to both *B* and *A*. Further addition of *C* will cause the state of the system to move along the solubility curve of *A*.

Exactly analogous phenomena will occur if we investigate the solubility of *C* in presence of increasing quantities of *B* in the solution. This is illustrated in Fig. 1, where the co-ordinates are the total concentrations of *B* and *C*, and the curves *A*, *B* and *C* represent the solubilities of these three solids. The points of intersection *AC* and *AB* are double points at which two solid phases are present simultaneously. (A third unstable double point occurs at the intersection of the *B* and *C* curves.)

Even when A -molecules are present at finite concentrations in the solution it is easily seen that the A -curve will be a hyperbola. Its axes are not however the co-ordinate axes, but are those represented by dotted lines in the figure, where the distance between the two origins is equal to the concentration of A -molecules in a solution saturated with A . This concentration is also equal to the increase of solubility along the straight lines A and B between pure solvent and the double points (marked with dotted lines in the figure).

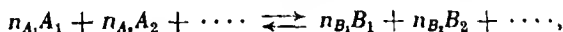
Curve systems of this type often occur in non-electrolyte systems (e.g., anthracene + picric acid, or naphthalene + picric acid) where the components combine to give a solid molecular compound consisting of one molecule of each component.

Equilibria corresponding to equation (25), in which the solid phase is a compound of the two components, are of particular importance in connection with electrolytes. The product of the concentrations of the components in the saturated solution is termed the *solubility product*. It is seen easily from the above considerations that the product of the concentrations of the dissociation products in the saturated solution must be a constant, both for the simple case treated above and for more complicated dissociations.

Systems resembling that represented in Fig. 1 will be dealt with further in Chapter VII.

4. THE EFFECT OF TEMPERATURE

(a) **Thermodynamic Treatment.** According to equation (8), the equilibrium of the mobile system



is determined by the equation

$$-\sum n_{A_i}\mu_{A_i} + \sum n_{B_i}\mu_{B_i} = 0,$$

the standard states being chosen in such a way that this equation is satisfied for a single possible state of the system. Introducing IV. (13)

$$\mu_{A_i} - \mu_{A_i(0)} = RT \ln p_{A_i}, \quad (27)$$

(where $\mu_{A_i(0)}$, the potential of A_i at unit pressure, depends only on the temperature), we obtain

$$\sum n_{A_i} \mu_{A_i(0)} - \sum n_{B_i} \mu_{B_i(0)} = RT \ln K_p. \quad (28)$$

We will introduce the following abbreviations for the sums occurring in this equation,

$$\sum n_{A_i} \mu_{A_i(0)} = \mu_{A(0)}, \quad (29)$$

$$\sum n_{B_i} \mu_{B_i(0)} = \mu_{B(0)}, \quad (30)$$

and shall use a similar terminology for the other thermodynamic functions. We thus obtain

$$\mu_{A(0)} - \mu_{B(0)} = RT \ln K_p, \quad (31)$$

Differentiating this equation with respect to temperature, remembering that

$$\left(\frac{\partial \mu_i}{\partial T} \right)_p = -S_i,$$

leads to the following equation,

$$RT \frac{d \ln K_p}{dT} + R \ln K_p = -(S_{A(0)} - S_{B(0)}), \quad (32)$$

or, introducing (31),

$$RT^2 \frac{d \ln K_p}{dT} + \mu_{A(0)} - \mu_{B(0)} = -T(S_{A(0)} - S_{B(0)}).$$

Using I. (30), $H = G + TS$, $H_1 = \mu_1 + TS$, we obtain finally

$$RT^2 \frac{d \ln K_p}{dT} = H_{B(0)} - H_{A(0)} = H_B - H_A = Q_p, \quad (33)$$

where Q_p is the heat absorbed by the process $A \rightarrow B$ at constant (but otherwise arbitrary) pressure.

Since $K_p = \frac{p_B}{p_A}$, this equation shows that an increase of temperature will favour the system having the greatest H -value, i.e., the system which is formed from the other system by absorption of heat. This agrees with Le Chatelier's principle.

An expression containing K_c instead of K_p can be derived from (33) by introducing (6),

$$K_p = K_c (RT)^{\sum n_{B_i} - \sum n_{A_i}}.$$

The expression thus obtained is

$$RT^2 \frac{d \ln K_c}{dT} = E_{B(0)} - E_{A(0)} = \Delta E = Q_v, \quad (34)$$

where Q_v is the heat absorbed by the process $A \rightarrow B$ at constant

volume. The variation of K_0 is thus related to the energy in the same way as the variation of K_p is related to the heat content.

Both (33) and (34) can also be easily derived by applying the second law in the form

$$q = T \left(\frac{dA}{dT} \right)_p,$$

to the complete process (1), using the expression in equation (4) for the work obtainable from the process.

By integrating the differential equation (33) on the assumption that Q is constant we obtain

$$\ln K_p = -\frac{Q}{RT} + I, \quad (35)$$

where I is a constant. This equation can also be written in the form

$$\ln \frac{K_{p,1}}{K_{p,2}} = -\frac{Q}{R} \left(\frac{1}{T_1} - \frac{1}{T_2} \right), \quad (36)$$

which shows how the equilibrium constant at one temperature can be calculated from its value at another temperature and the heat change associated with the process. (35) shows that there is a linear relation between $\ln K_p$ and $\frac{1}{T}$, which may be used for a simple graphical treatment of experimental data. However, as mentioned above, these expressions depend on the assumption that Q can be considered as independent of the temperature.

If this is not the case, Q may be represented by an empirical equation

$$Q = Q_0 + \alpha T + \beta T^2 + \gamma T^3 + \dots,$$

when integration of (33) leads to the following expression:—

$$-\ln K_p = \frac{Q_0}{RT} - \frac{\alpha}{R} \ln T - \frac{\beta}{R} T - \frac{\gamma}{2R} T^2 - \dots + I, \quad (37)$$

where Q_0 , α , β , ... are thermal quantities (i.e., quantities which can be determined from measurements of heat changes and specific heats), while I is a constant which cannot be derived from thermal data. The relation between α , β , ... and the specific heats is obtained from equation I. (71),

$$\left(\frac{dQ}{dT} \right)_p = c_{p(B)} - c_{p(A)},$$

since if the specific heats are given by the equations

$$\left. \begin{aligned} c_{p(A)} &= \alpha_1 + \beta_1 T + \gamma_1 T^2 + \dots \\ c_{p(B)} &= \alpha_2 + \beta_2 T + \gamma_2 T^2 + \dots \end{aligned} \right\} \quad (39)$$

it is easily shown that

$$\left. \begin{aligned} \alpha &= \alpha_2 - \alpha_1, \\ \beta &= \frac{1}{2}(\beta_2 - \beta_1), \\ \gamma &= \frac{1}{3}(\gamma_2 - \gamma_1), \end{aligned} \right\} \quad (40)$$

etc.

Analogous expressions are obtained by using K_c , ΔE and c_p in place of K_p , Q and c_p .

The above expressions for the effect of temperature on homogeneous equilibria are based on the first and second laws of thermodynamics, i.e., on "classical" thermodynamics. In a later section [X. 1.] we shall deal with certain results obtained in this field by considerations of molecular theory, in particular the so-called "third law of thermodynamics."

(b) **Gaseous Systems.** At constant temperature the nitrogen peroxide equilibrium



obeys the equilibrium expressions (14) and (17). By combining the first of these equations with (34) we obtain the following expression for the variation of the degree of dissociation with temperature at constant volume,

$$\ln \frac{\alpha_1^2 (1 - \alpha_2)}{\alpha_2^2 (1 - \alpha_1)} = \frac{\Delta E}{R} \left(\frac{1}{T_1} - \frac{1}{T_2} \right). \quad (41)$$

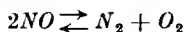
By inserting in this equation the degrees of dissociation at two temperatures (determined from the densities: cf. (18)), we find

$$\Delta E = 12,900 \text{ cal.}$$

in agreement with the value determined calorimetrically,

$$\Delta E = 12,500 \text{ cal.}$$

As for most other gaseous dissociations, the heat produced on dissociation is negative. The process of dissociation thus absorbs heat, and the degree of dissociation should increase with increasing temperature, in agreement with experiment. In contrast to this we may take the reaction



which is a "dissociation" from a chemical point of view, but which is not characterised by any increase in the number of molecules. The decomposition of nitric oxide is accompanied by the evolution of heat, and an increase of temperature thus renders the compound

more stable with respect to its components. The heat change in this reaction is practically independent of the temperature, since the molar specific heats of the three compounds are almost identical: hence the simple expression (35), or the analogous one,

$$\ln K_c = -\frac{\Delta E}{RT} + I, \quad (42)$$

can be used over a large temperature range. Table II shows the equilibrium concentration of nitric oxide in air at various temperatures.

TABLE II. *The Nitric Oxide Equilibrium*

T	% NO (exp.)	% NO (calc.)
1811	0.37	0.35
1877	0.42	0.43
2033	0.64	0.67
2195	0.97	0.98
2580	2.05	2.02
2675	2.23	2.35

The calculated percentages of *NO* are obtained from (42), using $\Delta E = 43,200$ cals. and taking for I the value giving the best agreement with experiment over the whole temperature range. It will be seen that nitric oxide is not very stable even at temperatures above $2,000^\circ$.

In using the above equations it must be remembered that \ln means natural logarithms, and that the gas constant $R = 1.985$. If decadic logarithms are used, R must be replaced by the factor

$$2.303R = 4.57.$$

A reaction which has been accurately studied over a large temperature range is the dissociation of iodine,



The relation between $\frac{1}{T}$ and $\log_{10} K_p$ for this reaction is shown in Fig. 2. It is almost linear, as would be expected from (35). In agreement with the negative heat of dissociation, the constant increases with increasing temperature and reaches a value of $K_p = 1$ at about $T = 1,450^\circ$. Application of equation (35) gives $Q = 34,500$ cals.

The dissociation of water vapour into oxygen and hydrogen has also been measured accurately at high temperatures. At temperatures of about $1,000^{\circ}$ the equilibrium can be investigated by a "streaming method," in which the water vapour is passed through a tube heated to a constant temperature so that equilibrium is set up during passage through the tube. If the equilibrium mixture

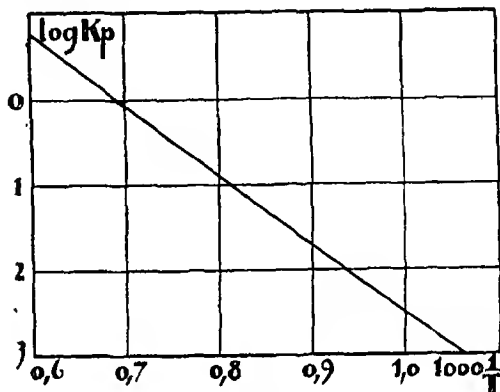
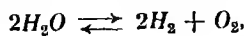


Fig. 2.

is rapidly cooled it is possible to "freeze" the equilibrium, so that no change of composition takes place during cooling. If the quantity of water vapour is known, the degree of dissociation can easily be calculated from analyses of the amounts of hydrogen and oxygen formed.

At higher temperatures, *e.g.*, $1,500^{\circ}$, it is not possible to cool the mixture sufficiently rapidly to avoid a displacement of the equilibrium while cooling, and this method is therefore not applicable in this temperature range. Use may be made of the fact that platinum and iridium are permeable to hydrogen at high temperatures, but not to oxygen or water vapour: by this means it is possible to measure directly the partial pressure of hydrogen in the equilibrium mixture. At still higher temperatures the degree of dissociation can be calculated from measurements of the pressure produced when mixtures of hydrogen and oxygen explode.

The relations between the concentrations, the total pressure and the degree of dissociation for the process



are easily seen to be

$$c_{H_2} = \frac{p}{RT} \cdot \frac{2\alpha}{2 + \alpha},$$

$$c_{O_2} = \frac{p}{RT} \cdot \frac{\alpha}{2 + \alpha},$$

$$c_{H_2O} = \frac{p}{RT} \cdot \frac{2 - 2\alpha}{2 + \alpha},$$

whence, introducing (6),

$$K_p = K_c RT = p \frac{\alpha^3}{(2 + \alpha)(1 - \alpha)^2}.$$

The variation of K_p with temperature can be calculated from equation (33), the thermal quantities used being the heat of combustion of hydrogen (57,650 cal. per mole at constant volume at 100°) and the specific heats of the substances concerned. Values of the degree of dissociation α can then be calculated from the last equation, and are given in the last column of Table III. They are compared with the values obtained experimentally by the methods given above. T is the absolute temperature, and the data refer to a pressure of one atmosphere.

TABLE III. *The Dissociation of Water Vapour*

T	$10^2\alpha$ (exp.)	$10^3\alpha$ (calc.)
1300	0.0027	0.0029
1397	0.0078	0.0084
1480	0.0189	0.0185
1500	0.0197	0.0221
1561	0.034	0.0368
2155	1.18	1.18
2257	1.77	1.79
2300	2.6	2.08

Water vapour is thus only about 2% dissociated even at 2,000° C.

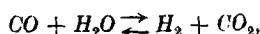
Similar considerations apply to the dissociation of carbon dioxide. At a pressure of one atmosphere the following figures are found for the degree of dissociation :—

TABLE IV. *The Dissociation of H_2O and CO_2*

T	$\alpha (H_2O)$	$\alpha (CO_2)$
1000*	$3.02 \cdot 10^{-7}$	$1.58 \cdot 10^{-7}$
1500	$1.97 \cdot 10^{-4}$	$4.06 \cdot 10^{-4}$
2000	$5.88 \cdot 10^{-3}$	$1.77 \cdot 10^{-2}$
2550	$3.98 \cdot 10^{-1}$	$1.58 \cdot 10^{-1}$

It will be seen that the dissociation of carbon dioxide increases more rapidly with temperature than the dissociation of water vapour: this is in agreement with the fact that the heat of combustion of carbon monoxide is greater than that of hydrogen.

The equilibrium constants for the dissociation of water vapour and carbon dioxide can be used to calculate the equilibrium in the *water gas reaction*,



for which

$$\frac{c_{H_2} c_{CO_2}}{c_{CO} c_{H_2O}} = K.$$

writing $\frac{c_{H_2}^2 c_{O_2}}{c_{H_2O}^3} = K_1,$ $\frac{c_{CO}^2 c_{O_2}}{c_{CO_2}^2} = K_2,$ (43)

we find

$$K = \sqrt{\frac{K_1}{K_2}}.$$

Determinations of K at different temperatures have been used to verify the dissociation constants found for water vapour and carbon dioxide.

Equation (43) is an example of the general rule that if a reaction can be written as the sum of several partial reactions, then its equilibrium constant can be expressed as the product of the equilibrium constants for the partial reactions.

The equilibrium between ammonia and its constituent elements has been specially investigated in the temperature range 300° – 500° C. in presence of catalysts. The dissociation constant

$$K_p = \frac{P_{H_2}^3 P_{N_2}}{P_{NH_3}^2}$$

increases rapidly with increasing temperature. Increase of pressure will clearly favour the formation of ammonia: this fact is of importance in the industrial synthesis of ammonia. At very high pressures

K_p decreases, i.e., the yield of ammonia is greater than the value calculated from the gas laws.

Table V gives some values for the percentage of ammonia in an equivalent mixture of the components when equilibrium has been reached. t is the Centigrade temperature, and p the pressure in atmospheres.

TABLE V. *The Ammonia Equilibrium*

$t \backslash p$	10	100	1000
350	7.35	—	—
400	3.85	24.91	—
450	2.04	16.36	69.4
500	1.20	10.40	—

(c) **Homo-heterogeneous Systems.** The temperature dependence of the equilibrium constant of a homogeneous reaction is given by (33),

$$RT^2 \frac{d \ln K_p}{dT} = Q_p.$$

If this equation is combined with equation II (49) for the heat of vaporisation of the substance A_1 (e.g., in the solid state),

$$RT^2 \frac{d \ln p_{A_1}}{dT} = q_{A_1},$$

we obtain

$$RT^2 \frac{d \ln K'_p}{dT} = Q'_p, \quad (44)$$

where

$$K'_p = K^p p_{A_1}^{n_{A_1}}, \quad (45)$$

and

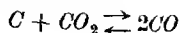
$$Q'_p = Q_p + n_{A_1} q_{A_1}. \quad (46)$$

K'_p is obviously the mass action constant for the homo-heterogeneous reaction $A \rightarrow B$ in which the substance A_1 takes part as a solid (cf. [3. a.]). Further, it is clear that Q'_p is the heat change for the same process.

This result can be expressed in a general form as follows. Let one or more of the components in the equilibrium scheme (1) react in the solid form, and let the mass action constants obtained by

omitting the factors corresponding to these components be K'_p and K'_c . Then the variation of K'_p and K'_c with temperature is given by equations analogous to (33) and (34), using the heat change for the reaction (1) when the components in question react as solids.

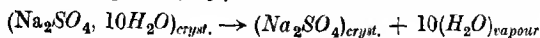
For example, for the equilibrium



the temperature variation of the constant $K_p = \frac{P_{CO}^2}{P_{CO_2}}$ is obtained

from equation (44) by introducing the heat change for the reaction between solid carbon and gaseous carbon dioxide.

Specially simple relations are obtained when the gas phase only contains one component, *e.g.*, in the reaction



we have, since $K_p = P_{H_2O}$:

$$RT^2 \frac{d \ln P_{H_2O}}{dT} = Q,$$

where Q is the heat absorbed in the efflorescence of the hydrated salt $(Na_2SO_4, 10H_2O)_{cryst.}$ according to the above equation.

CHAPTER VII

ELECTROLYTES

1. THE THEORY OF ELECTROLYTIC DISSOCIATION

MOST substances which are termed salts (and to a certain extent acids and bases also) exhibit an "electrolytic conductivity" in aqueous solution. When a current passes through their solutions, matter is transported. These substances also exhibit anomalies in their osmotic behaviour, *e.g.*, the freezing point depression and vapour pressure lowering of their solutions are abnormally large. It was on observations of this kind that Arrhenius based his theory of electrolytic dissociation (1884-87).

According to this theory, the molecules of substances possessing the above characteristics are dissociated in aqueous solution into electrically charged components, known as ions. Thus, for example, potassium chloride dissociates according to the scheme $KCl \rightarrow K^+ + Cl^-$. The electrical conductivity of the solution is then due to the motion of the ions in the electric field, and the increase in the freezing point depression to the increase in the number of solute molecules.

The classical electrolytic dissociation theory supposed that a dissociation equilibrium was set up in solution, *e.g.*, for potassium chloride



If the degree of dissociation is α , the number of solute molecules formed from one molecule of KCl will obviously be $1 + \alpha$. Provided, that the laws of dilute solutions can be applied to the mixture, the factor $1 + \alpha$ will also represent the ratio of the observed molar freezing point depression $-\Delta$ to the normal freezing point depression $-\Delta_0$. In this case we have

$$\frac{\Delta}{\Delta_0} = 1 + \alpha = i, \quad (1)$$

where i is the so-called van't Hoff factor. Both i and α can be determined from measurements of the molar freezing point lowering, using for $-\Delta_0$ the value 1.86 previously given. The fact that $-\Delta$ is found to increase with increasing dilution is in agreement with

this interpretation, since according to the laws for dissociation equilibria [VI. 2. c.] α should increase in the same way.

The electrolytic conductivity can be used similarly for determining the degree of dissociation. Since the conductivity of one gram-molecule of salt is due entirely to the ions formed by dissociation, according to the classical theory it should be proportional in dilute solutions to the number of ions, *i.e.*, to the degree of dissociation. We should therefore expect that with increasing dilution the molar conductivity would increase, finally reaching a maximum value corresponding to complete dissociation. If this maximum conductivity is Λ_{∞} , and the conductivity corresponding to a degree of dissociation α is Λ , we have

$$\alpha = \frac{\Lambda}{\Lambda_{\infty}}, \quad (2)$$

We thus have two independent methods of determining α , and for a large number of salts there is reasonable agreement between the values found by the two methods. Arrhenius' theory is in general capable of accounting for a large proportion of the characteristic properties of electrolytes.

The equilibrium laws valid for dilute solutions lead to the mass action equation VI. (14),

$$c \frac{\alpha^2}{1 - \alpha} = K, \quad (3)$$

If, however, the values of α obtained by the methods described above are inserted in this equation, *no verification* is observed in the case of salts which are dissociated to a large extent, the quantity K exhibiting marked variations with concentration. Equations (1) and (3) are based on exactly the same assumptions, *i.e.*, the validity of the simple gas laws and the equation of state

$$PV = RT(1 + a) = RTi$$

derived from them. It is therefore a thermodynamic necessity that *the simple gas laws are not applicable to strongly dissociated electrolytes* and that the degrees of dissociation calculated on the above assumptions must be considered as illusory. The same conclusion is reached by considering a number of other cases in which solutions of this kind deviate from the simple laws applying to non-electrolytes. All these deviations are referred to together as "the anomalies of strong electrolytes."

In the case of strongly dissociated substances these anomalies are present even in very dilute solutions, and it is natural to attribute **them** to the presence of electric charges on the ions. On this basis

it has proved possible to develop a theory of electrolytes which is in most respects satisfactory.

2. STRONG ELECTROLYTES

(a) **The Complete Dissociation of Strong Electrolytes.** While properties such as the electrolytic conductivity and the freezing point appear to indicate incomplete dissociation of salts like potassium chloride and other strong electrolytes, there are other properties *e.g.*, the colour or catalytic activity, which indicate that these electrolytes are completely dissociated into ions at all ordinary concentrations. The modern theory of strong electrolytes assumes that the latter interpretation corresponds to the true state of affairs, and that the conclusions previously drawn from freezing points and conductivities are untenable because the electric forces between separated ions were not taken into account (Bjerrum).

This conception harmonises with the previously mentioned fact that substances like *KCl* and *NaCl* form ionic lattices in the solid state, and are not present as salt molecules.

It is clear that there must be forces of attraction between oppositely charged ions in an electrolyte solution. These forces will tend to bring the ions together with the formation of ion pairs held together by electrostatic forces. On the other hand, thermal agitation will tend to dissociate such "electrostatic molecules" into their constituent ions. The net result of these two opposed tendencies will depend upon the temperature and the solvent, as well as on the ions themselves. An increase of temperature will increase the tendency to separation arising from the molecular motion, and the same result will be produced by passing from a solvent of low dielectric constant to one of high dielectric constant, since the electrostatic forces between two charged particles are inversely proportional to the dielectric constant of the surrounding medium. Water has a very high dielectric constant, with the result that typical binary electrolytes in aqueous solution can be assumed to consist essentially of free ions under ordinary experimental conditions. On the other hand, the electrostatic forces are large enough to affect both the thermodynamic properties and the conductivities of the ions, thus producing the illusion of incomplete dissociation in the classical treatment.

(b) **Activities and Activity Coefficients.** The most important problem in the theory of salt solutions is to determine the variation of the thermodynamic functions with the concentration of the solution. The activity function is particularly convenient for representing the effect of concentration at constant temperature.

For a binary electrolyte like *KCl* the thermodynamic law of mass

action VI. (23) will hold, independent of the degree of dissociation of the electrolyte, *i.e.*,

$$a_{K^+} a_{Cl^-} = K a_{KCl}, \quad (4)$$

where K is a constant depending on the temperature but not on the concentration. An equation of this kind is purely thermodynamic in character, and is valid independent of whether or not the molecules in question are actually present in the solution. If however (in agreement with modern views) the concentration of undissociated salt molecules is practically zero and cannot be determined by any means, then the insertion of any arbitrary numerical value for this concentration will have no effect on the thermodynamic treatment of the salt solution. It is therefore permissible, and particularly convenient, to standardise the activities in such a way that equation (4) assumes the simplest possible form, *i.e.*, putting $K = 1$. We then obtain a purely conventional definition of the activity of the salt in terms of the activities of the ions,

$$a_{K^+} a_{Cl^-} = a_{KCl}, \quad (5)$$

or more generally, for any binary salt,

$$a_+ a_- = a_{\pm}^2 = a_s. \quad (6)$$

where a_s is the activity of the salt and a_{\pm} the "mean activity" of the ions.

It has been previously mentioned [V. 1] that it is advantageous to choose the arbitrary proportionality factor occurring in the definition of activity in such a way that the activity and mole fraction x of the solute became identical at infinite dilution. In the present instance we can standardise the activity of the ions, when the activity of the salt will also be fixed according to the above equations. Instead of defining the activity in terms of the mole fraction, as above, it is more convenient to standardise the activity of the ions so that it coincides with the ionic molar concentration (or molarity) at infinite dilution. If we use the symbol m to denote the molar concentration of both ions and salts, then for a binary electrolyte at infinite dilution we have

$$a_{\pm} = a_+ = a_- = m = m_+ = m_-. \quad (7)$$

In conformity with this, the activity coefficient of a solute species will be defined as the ratio between the activity and the molar concentration of that species, so that we can write

$$f_+ = \frac{a_+}{m_+}, \quad f_- = \frac{a_-}{m_-}, \quad (8)$$

while the mean ionic activity coefficient and the activity coefficient of the salt are correspondingly defined by the simple equation

$$f_+ f_- = f_{\pm}^2 = f_s. \quad (9)$$

For a salt of higher valency type the corresponding expressions are of course somewhat more complicated. If the salt dissociates into ν_+ cations and ν_- anions, where $\nu_+ + \nu_- = \nu$, then

$$m_+ = m\nu_+, \quad m_- = m\nu_-,$$

and

$$a_+^{\nu_+} a_-^{\nu_-} = a_{\pm}^{\nu} = a_s,$$

whence at infinite dilution,

$$a_+ = m_+ = m\nu_+,$$

$$a_- = m_- = m\nu_-,$$

and

$$a_{\pm}^{\nu} = a_s = m_+^{\nu_+} m_-^{\nu_-} = m^{\nu} \nu_+^{\nu_+} \nu_-^{\nu_-}. \quad (10)$$

On the basis of (10) we can define the activity coefficients as

$$f_s = \frac{a_s}{m^{\nu} \nu_+^{\nu_+} \nu_-^{\nu_-}}, \quad (11)$$

and

$$f_{\pm} = \frac{a_{\pm}}{\nu_+^{\nu_+} \nu_-^{\nu_-} m^{\nu}}, \quad (12)$$

or by an equation analogous to (9),

$$f_+^{\nu_+} f_-^{\nu_-} = f_s^{\nu} = f_{\pm}^{\nu}, \quad (13)$$

which makes coefficients assume the value unity at infinite dilution.

If in these expressions we put $\nu_+ = \nu_- = 1$, we obtain the simple expressions (6) and (9) for the activity and activity coefficients of a uni-univalent salt.

It should however be noted that the activity and activity coefficient of a single ionic species are not accessible by ordinary thermodynamic methods. These methods depend on considerations of the work obtainable by transferring the molecular species in question from one state to another, and when dealing with charged particles this work of transfer depends to a very great extent on the distribution of electric charges in the phase under consideration. This problem does not arise as a rule, since the operations usually dealt with consist either of the simultaneous transfer of ions of opposite signs, or the replacement of one ion by another of the same charge: in either case the net transport of electricity is zero. The effect of the charge can be eliminated in principle only if the phases between which the transfer takes place are electrically neutral.

The thermodynamic behaviour of a salt solution at constant temperature can be completely described in terms of activity coefficients. In the following sections we shall deal with the most important methods available for determining the activity coefficients of salts.

(c) **The Variation of Activity Coefficients with Concentration.** The effects due to the electric charges of the ions are apparent even in the dilute range in which non-electrolytes usually obey the simple gas laws with considerable accuracy. Since the deviations from ideal behaviour are due to the charges, we should expect that dilute salt solutions of the same charge type will behave in approximately the same way.

It has already been pointed out that the ions in a salt solution are not distributed in a perfectly random manner, but that positive ions will tend to be surrounded by an excess of negative ions, and *vice versa*. Each ion will be surrounded by an "ionic atmosphere" in which there is an excess of ions of opposite sign. This arrangement of ions will diminish their chemical potential, and the effect will increase with increasing salt concentration. Hence on account of the relation IV. (61) between chemical potential and activity coefficient, the latter will decrease with increasing salt concentration.

These ideas have been developed quantitatively by methods which will be described later. For a uni-univalent salt like KCl dissolved in water at ordinary temperatures it is found that the mean activity coefficient is given by the equation

$$-\log f_{\pm} = \alpha \sqrt{m}, \quad (14)$$

or, if we assume that the activity coefficients of the two oppositely charged ions are both equal to f_1 ,

$$-\log f_1 = \alpha \sqrt{m}, \quad (15)$$

where α is a constant having a value of almost exactly 0.5 at ordinary temperatures. These equations are however only valid for very small concentrations, considerably below 0.01 molar. For concentrations up to about 0.1 molar the activity coefficient appears to be given by adding a linear term to (15), giving

$$-\log f_1 = -\log f_{\pm} = \alpha \sqrt{m} - \beta m, \quad (16)$$

where β depends on the nature of the salt. For the alkali chlorides β is about 0.5.

It will be seen from these equations that the curve representing

the relation between the activity coefficient and the concentration must have an infinite slope at $m = 0$, since

$$\left(\frac{d \ln f_1}{dm}\right)_{m=0} = \left(\frac{df_1}{dm}\right)_{m=0} = -\infty. \quad (17)$$

It also follows from (16) that there is a linear relation between $\ln f_1$ and \sqrt{m} in the dilute range. This is illustrated in Fig. 1, which

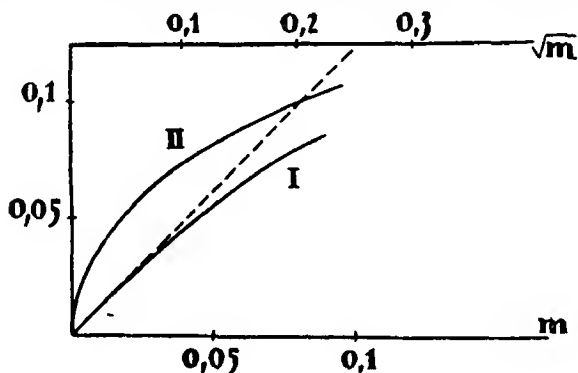


FIG. 1.

represents the behaviour of potassium chloride. The ordinate is $-\log_{10} f_1$, while the abscissæ for curves I and II are \sqrt{m} and m respectively.

Fig. 2 shows the variation of f with m , also for KCl , while Table I gives the corresponding values of f_{\pm} at round concentrations.

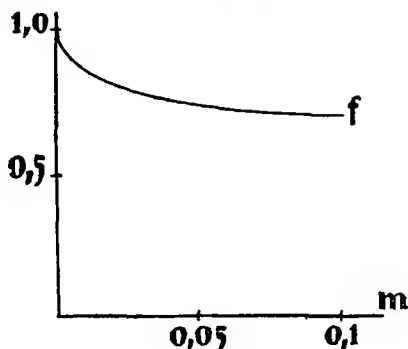


FIG. 2.

TABLE I. *Activity Coefficients of KCl*

m	f_{\pm}
0.0001	0.989
0.001	0.966
0.01	0.902
0.1	0.780

At high dilutions the individual properties of different salts are not in evidence, and in agreement with (15) the activity coefficients of ions of the same charge will be independent of the nature of the salt. On the other hand, we should expect that the activity coefficient will depend on the charge of the ion considered, and also on the charges of the other ions present in the solution. The effect of this factor is expressed by the equation

$$-\log f_z = \alpha z^2 \sqrt{\mu}, \quad (18)$$

where α has the value previously given, f_z is the activity coefficient of an ion of charge (or valency) z , and μ is the ionic strength. The ionic strength is defined by

$$\mu = \frac{1}{2} \sum m_i z_i^2 = \frac{1}{2} \sum c_i z_i^2, \quad (19)$$

where c is the equivalent concentration of any ion in the solution. For example, for a 0.01 molar $MgCl_2$ solution (19) gives

$$\mu = \frac{1}{2}(0.01 \cdot 2^2 + 0.02) = 0.03$$

or

$$\mu = \frac{1}{2}(0.02 \cdot 2 + 0.02) = 0.03$$

According to equation (18) the activity coefficient of an ion will thus be the same in solutions of the same ionic strength, *e.g.*, in 0.03 molar KCl and in 0.01 molar $MgCl_2$ or K_2SO_4 . However, in cases involving ions of high valency the equations given are only valid at high dilutions.

If ions of small and varying concentrations are dissolved in a salt solution of large and constant concentration, then the activity coefficients of the former ions will be practically constant, since for each ion the activity coefficient depends only on the total ionic concentration. The simple laws of dilute solutions are therefore applicable to such solutions.

It has previously been shown [Δ 1] that the laws of dilute solutions can be expressed in terms of the activity coefficients f_1 and f_2 .

for the components K_1 and K_0 , by the following equations, valid at infinite dilution,

$$x \frac{d \ln f_1}{dx} = 0, \quad \text{V. (13)}$$

$$\frac{d \ln f_0}{dx} = 0, \quad \text{V. (14)}$$

and it was also shown that Raoult's and Henry's laws will hold provided that these equations are satisfied. Using m in place of x , V. (13) can be rewritten as

$$m \frac{d \ln f_1}{dm} = 0. \quad (20)$$

In order to investigate whether these equations are also valid for electrolytes, the expression (16) may be used. On differentiating with respect to m , it gives

$$-\frac{d \log f_1}{dm} = \frac{1}{2} a m^{-\frac{1}{2}} - \beta,$$

or

$$-0.4343 m \frac{d \ln f_1}{dm} = \frac{1}{2} a m^{\frac{1}{2}} - \beta m.$$

Since the right-hand side of this equation becomes zero for $m = 0$, equation (20) is satisfied. Henry's law will thus hold as a limiting law for the individual molecules or ionic species, in spite of the fact that from (17)

$$\frac{df_1}{dm} = -\infty,$$

at $m = 0$. It is easily seen that the same will apply to salt solutions of higher valency types.

Since V. (13) and V. (14) are interdependent, Raoult's law must also hold as a limiting law for salt solutions.

Instead of expressing the laws of electrolyte solutions in terms of the activity coefficients of the ions, we may also use either the activity coefficient of the solute, f_0 or the *osmotic coefficient* ϕ . The relation between f_0 and ϕ is obtained in the following way. For ideal solutions we have from IV. (57),

$$\mu_0 - \mu_{0(0)} = RT \ln(1 - x), \quad (21)$$

while for non-ideal solutions the concentration of the solvent must be replaced by the activity $a_0 = (1 - x)f_0$. We therefore have in general,

$$\mu_0 - \mu_{0(0)} = RT \ln(1 - x)f_0. \quad (22)$$

Further, V. (6) gives as a limiting law for ideal solutions,

$$\mu_0 - \mu_{0(0)} = -RTx, \quad (23)$$

while we can write in general

$$\mu_0 - \mu_{0(0)} = -RTx\varphi, \quad (24)$$

the last equation constituting a definition of the osmotic coefficient φ . When these equations are applied to salts, x is the sum of the mole fractions of the ions.

It is easily seen that the osmotic coefficient for salts is related to i by the following general equation,

$$\varphi = \frac{i}{\nu}, \quad (25)$$

where i is the van't Hoff factor determined thermodynamically and ν the number of ions formed when one molecule of salt dissociates completely.

(22) and (24) give

$$\ln(1-x)f_0 = -\varphi x, \quad (26)$$

or, putting

$$\ln(1-x) = -x$$

(which is permissible for small values of x),

$$\ln f_0 = (1-\varphi)x \quad (27)$$

which is the required relation between f_0 and φ .

It is now possible to derive a relation between the osmotic coefficient φ and the activity coefficient of the salt, assuming the solution to contain a single salt the concentration of which can vary. This relation is obtained by using the Gibbs-Duhem equation in the form V. (15),

$$x \frac{d \ln f_1}{dx} + (1-x) \frac{d \ln f_0}{dx} = 0.$$

If this is introduced in equation (27) differentiated,

$$\frac{d \ln f_0}{dx} = (1-\varphi) - x \frac{d\varphi}{dx}, \quad (28)$$

we obtain

$$x \frac{d \ln f_1}{dx} = x \frac{d\varphi}{dx} - (1-\varphi), \quad (29)$$

If the component K_1 is a salt, it is easily seen that the mean activity coefficient f_{\pm} must be used in place of f_1 . Since further there is

direct proportionality between m and x , (29) can be rewritten as

$$m \frac{d \ln f_{\pm}}{dm} = m \frac{d\varphi}{dm} - (1 - \varphi), \quad (30)$$

where m is the molality of the salt. It will be seen that a knowledge of φ at a given concentration is not sufficient for calculating f , but that it is necessary to integrate over the range zero to m , and hence to know φ over the whole of this range.

We can apply equation (30) to solutions of a binary electrolyte which obeys equation (16). Using (30), it is easily seen that a relation of the form

$$1 - \varphi = a \sqrt{m} + bm,$$

leads to

$$- \ln f_{\pm} = 3a \sqrt{m} + 2bm,$$

whence for sufficiently dilute solutions,

$$d \ln f_{\pm} = 3d\varphi, \quad (31)$$

If therefore (16) is assumed to represent f , the osmotic coefficient must be given by

$$0.4343 (1 - \varphi) = \frac{\alpha}{3} \sqrt{m} - \frac{\beta}{2} m. \quad (32)$$

The coefficients α and β can thus be determined by measuring the osmotic coefficient of the solution over a sufficiently large concentration interval.

Another expression for f has been proposed, which is applicable at somewhat higher concentrations than (16), but which does not agree with experiment in the most dilute range. It is

$$- \log f = \alpha \sqrt[3]{m}, \quad (33)$$

where α is a constant characteristic of the salt.

We shall now deal with a few of the most important methods which have been used for determining the activity coefficients of strong electrolytes.

(d) **The Solubility of Strong Electrolytes.** If determinations are made of the solubility of a sparingly soluble salt in aqueous solutions of a second salt having no ion in common with the first, it will be found that the solubility increases with the concentration of the second salt. An example of this is shown in Fig. 3, where the solubility of $AgCl$ in sodium nitrate solutions is plotted against the concentration of the sodium nitrate.

According to the classical theory of electrolytes, the "solubility product" [VI. 3. c.] is constant, and the increase in solubility observed was attributed to the formation of undissociated salt-molecules in the solution. Thus in the example considered here

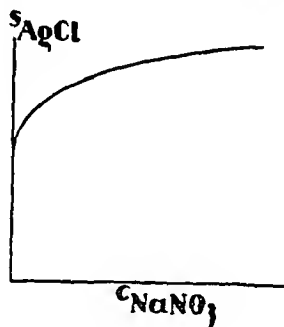


FIG. 3.

the reaction $\text{AgCl} + \text{NaNO}_3 \rightarrow \text{AgNO}_3 + \text{NaCl}$ was supposed to take place, the increase in the silver content of the solution being identified with its content of AgNO_3 molecules.

However, on the basis of modern ideas about the dissociation of salts (as developed above), the increase in solubility must be attributed to other causes. Since the gas laws are not valid for the ions, the thermodynamic law of mass action and [VI. 2. e.] show that the law for the "solubility product" of AgCl must be written in the form

$$a_{\text{Ag}} + a_{\text{Cl}^-} = K_a, \quad (34)$$

where a is the activity and K_a a constant independent of the concentration which is termed the *thermodynamic solubility product*. Equation (34) can be written

$$m_{\text{Ag}} + f_{\text{Ag}} + m_{\text{Cl}^-} f_{\text{Cl}^-} = K_a, \quad (35)$$

or

$$K_c f_1^2 = s^2 f_1^2 = K_a, \quad (36)$$

or

$$\log s + \log f_1 = \frac{1}{2} \log K_a. \quad (37)$$

where K_c is the ordinary solubility product, s the solubility and f_1 the activity coefficient of a univalent ion. If the salt in the solvent

is a uni-univalent salt like NaNO_3 , and the solution is sufficiently dilute for (14) and (15) to be applicable, then (37) gives

$$\log s = \frac{1}{2} \log K_a + \alpha \sqrt{m}, \quad (38)$$

where m is the total molarity of the solution. (38) can also be written as

$$\frac{d \log s}{d \sqrt{m}} = \alpha.$$

It is thus possible to determine the constant α in equations (14) and (15) from solubility measurements. If the observed values of $\log_{10} s$ are plotted against \sqrt{m} , then in the range where (15) is valid a straight line of slope α will be obtained, while the intercept on the solubility axis will be equal to $\frac{1}{2} \log K_a$.

If the solubility of the salt in pure water is s_0 , and if the value of this is so small that the activity coefficients of the ions can be taken as unity, then from (36) we can write $K_a = s_0^2$, so that (38) becomes

$$\log \frac{s}{s_0} = \alpha \sqrt{m}. \quad (39)$$

Experiments of this kind have been used to determine the value of α . At higher salt concentrations equation (16) can be applied in the same way, but in this case the coefficients β will depend in an individual way on the nature of the electrolytes concerned.

It is obvious that the solubilities of salts of higher valency types can be used in the same way, though in this case deviations from the simple formulae are encountered even at very low concentrations.

If the sparingly soluble salt has an ion in common with the salt which is used as solvent (*e.g.*, AgCl in a KCl solution), the solubility product will vary with the total salt concentration in the same way as in the preceding case. Equation (35) is therefore again valid, but in this case $m\text{Cl}$ is not equal to s , but to $m + s$, where m is the molar concentration of the KCl solution. We therefore have in place of (36),

$$s(m + s)f_1^2 = K_c f_1^2 = K_a, \quad (40)$$

or

$$\log s + \log(m + s) + 2 \log f_1 = \log K_a, \quad (41)$$

or

$$\log s + \log(m + s) = \log K_a + 2\alpha \sqrt{m + s}, \quad (42)$$

where again α can be determined from the solubility values.

It follows from (42) that the addition of a salt having a common ion has a dual effect upon the solubility. On the one hand it diminishes the solubility on account of the term $\log(m + s)$, while on the other hand it increases the solubility on account of the term $2\alpha\sqrt{m + s}$. For the simple types of salts we are dealing with here the first effect will be predominant, but on account of the opposite effect expressed by the square root term the solubility lowering will be smaller than it would be in the absence of electric forces between the ions.

If more concentrated salt solutions are used as solvent it is no longer possible to give any simple rules for predicting the effect of salt concentration upon the solubility, since the individual factors become of decisive importance. In certain cases the increase of solubility caused by addition of salts in very dilute solutions will

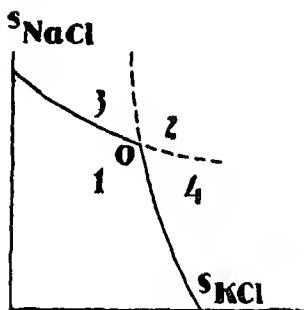


FIG. 4.

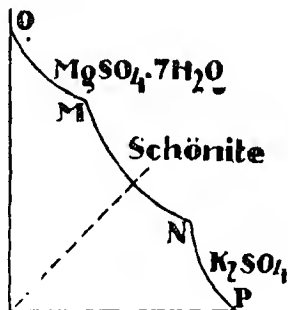


FIG. 5.

continue into the concentrated range, and very large effects can be produced, particularly with solutes of high valency types. In other cases the "salting out effect" (i.e., the decrease in solubility which salts generally exert on non-electrolytes) will come into action, in which case the solubility curve may pass through a maximum at a certain salt concentration.

The solubility diagram for a system of two or more electrolyte components can be treated from the point of view of the phase rule just as in the case of non-electrolytes. If the two salts constituting the system have a common ion, there will in general be a mutual solubility lowering, even if this lowering does not follow the simple laws valid for dilute solutions. The accompanying figures show three typical examples, the co-ordinates in each case being the stoichiometric concentrations of the two salts. Fig. 4 shows the behaviour of the system NaCl — KCl . The regions 1 and 2 are

respectively unsaturated with respect to both components, while region 3 is unsaturated with respect to KCl and supersaturated with respect to $NaCl$, and region 4 is supersaturated with respect to KCl and unsaturated with respect to $NaCl$. The dotted parts of the curves thus represent unstable states. Fig. 5 shows the diagram for the system K_2SO_4 — $MgSO_4$, where the solid phases are K_2SO_4 , $MgSO_4 \cdot 7H_2O$ and the double salt schönite, $K_2SO_4 \cdot MgSO_4 \cdot 6H_2O$. At N and M the double salt is in equilibrium with K_2SO_4 and $MgSO_4 \cdot 7H_2O$ respectively. Finally Fig. 6 illustrates the system KCl — $MgCl_2$, in which the double salt carnallite, $KCl \cdot MgCl_2 \cdot 6H_2O$, appears. It will be seen from the figures that the difference between

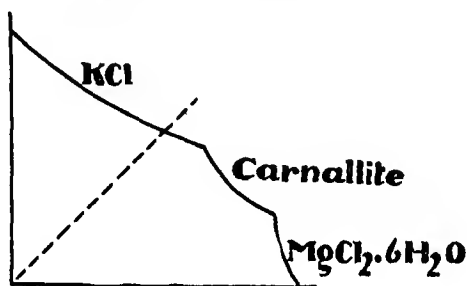


FIG. 6.

the last two figures lies in the fact that in the case of schönite the double salt can form a stable saturated solution on the addition of water, while in the case of carnallite the corresponding saturated solution is unstable and will spontaneously deposit solid KCl .

(e) **Determination of Activity Coefficients from Electromotive Force Measurements.** It will be seen in the chapter on electrochemistry that the electrical energy produced by a galvanic cell working reversibly can be taken as a measure of the work obtainable from the process producing the current. For example, if HCl is formed in a certain type of cell, then if we consider two cells of this type with different HCl concentrations the difference in the amounts of electrical energy produced by the two cells will determine the work of transfer of HCl , and hence from I. (53) and IV. (59) the difference in chemical potential and the ratio of the activities.

If the HCl concentrations are m_1 and m_2 and the electrically measured work of transfer per mole of HCl is A , then

$$A = 2RT \ln \frac{m_1 f_1}{m_2 f_2}, \quad (43)$$

which can be used to determine the relation between m and f .

(f) **The Freezing Points of Salt Solutions.** It has already been mentioned that the abnormal freezing point depressions of strong electrolytes were an important factor in the foundation of the theory of electrolytic dissociation. This is illustrated by the data in Tables II and III, which give the molar freezing point depressions of KCl and $BaCl_2$ at different concentrations: since the normal molar freezing point depression is 1.86, the abnormality is obvious.

TABLE II.— $-\Delta T$ for Potassium Chloride

m	$-\Delta T$ mol.	φ (exp.)	φ (calc.)
0.00025	3.71	0.997	0.994
0.002	3.658	0.983	0.984
0.005	3.643	0.980	0.975
0.01	3.609	0.968	0.967
0.02	3.570	0.959	0.957
0.05	3.505	0.942	0.941
0.10	3.451	0.927	0.934

TABLE III.— $-\Delta T$ for Barium Chloride

m	$-\Delta T$ mol.	φ (exp.)
0.0002	5.56	0.996
0.001	5.34	0.957
0.005	5.12	0.916
0.010	5.03	0.902
0.050	4.80	0.860
0.100	4.71	0.843

As previously mentioned, this behaviour is most readily expressed in terms of the osmotic coefficient φ , defined as

$$\varphi = \frac{i}{\nu} \quad (44)$$

where i is the van't Hoff factor, determined thermodynamically, and ν is the number of ions formed by the complete dissociation of a salt molecule. If the gas laws are obeyed and there is complete dissociation, φ is unity; hence its deviations from unity constitute a measure of the effect of the interionic forces.

As described in section (c), a knowledge of φ in the concentration

range 0 to m can be used to calculate f over the same range. Since freezing point determinations can be carried out with high accuracy, this method is of great importance for determining activity coefficients.

It should however be mentioned that data obtained from freezing point estimations do not strictly refer to a constant temperature. For more concentrated solutions a correction must be applied to allow for the complications thus introduced.

(g) **Laws Governing Volume and Thermal Qualities.** By applying the general expressions IV. (42) and IV. (45) to the volume of solutions, we obtain

$$V = xV_1 + (1-x)V_2, \quad (45)$$

$$x dV_1 + (1-x) dV_2 = 0, \quad (46)$$

and these equations are applicable both to electrolytes and to non-electrolytes.

When $x = 0$,

$$\frac{dV_1}{dV_2} = -\infty.$$

This condition is satisfied in a way which differs from the corresponding condition governing chemical potentials. In the case of non-electrolytes $\frac{dV_1}{dx}$ is finite and $\frac{dV_2}{dx}$ is zero, while in the case of electrolytes V_1 depends on the concentration in the same way as the activity coefficients of ions. It is thus seen that addition of solvent to a very dilute solution causes no change in total volume.

V_1 , the differential molar volume of a strong electrolyte in solution, is in general considerably less than the molar volume of the electrolyte in the pure state, and in many cases is even negative. This is due to the electric field surrounding the ions exerting a compressing effect on the water molecules, which have a marked electrical structure, though no net charge.

Of the thermal phenomena associated with electrolytes, the heat of dilution is most easily treated from a theoretical point of view. If we differentiate with respect to temperature equation (22),

$$\mu_0 - \mu_{0(0)} = RT \ln(1-x)f_0,$$

we obtain

$$\frac{d\mu_0}{dT} - \frac{d\mu_{0(0)}}{dT} = RT \frac{d \ln f_0}{dT} + R \ln(1-x)f_0, \quad (47)$$

or

$$-H_0 + H_{0(0)} = RT^2 \frac{d \ln f_0}{dT}. \quad (48)$$

It is thus seen that the heat of dilution is zero if f_0 is independent of the temperature. If the temperature variation of f_0 is known, (48) can be used to calculate $Q: H_0 - H_{0(0)}$, which is the differential heat of dilution,

i.e., the heat absorbed when one mole of K_0 is added to a large quantity of the solution. By means of equation (27), (48) can be expressed in terms of the osmotic coefficient instead of f_0 , giving

$$Q = RT^2 x \frac{d\varphi}{dT}. \quad (49)$$

The variation of the activity coefficients with temperature depends on two effects. In the first place an increase of temperature will produce an increase in f_{\pm} , since an increased thermal motion will decrease the effect of the electric forces. In the second place the temperature will affect the dielectric constant of the medium and hence the electric forces, which decrease with increasing D . It will be shown later that in sufficiently dilute solutions of a uni-univalent salt we have

$$-\ln f_{\pm} = \frac{B\sqrt{m}}{(DT)^{\frac{1}{2}}}, \quad (50)$$

where B is a positive constant. Differentiating (50) and introducing (31) we obtain

$$3 \frac{d\varphi}{dT} = \frac{d\ln f_{\pm}}{dT} = \frac{3}{2} BD^{-\frac{1}{2}} T^{-\frac{1}{2}} \left(1 + \frac{T}{D} \frac{dD}{dT} \right) \sqrt{m}, \quad (51)$$

or by introducing (49)

$$Q = H_0 - H_{0(w)} = \frac{xBR\sqrt{m}}{2D^{\frac{1}{2}}T^{\frac{1}{2}}} \left(1 + \frac{d\ln D}{d\ln T} \right). \quad (52)$$

Measurements of the dielectric constant of water show that in this case

$$\frac{d\ln D}{d\ln T} = -1.5,$$

and Q is hence negative. This means that heat is evolved when salt solutions are diluted. These calculations are however only valid for extremely dilute solutions.

For salts such as KCl , $NaNO_3$, etc., the *heat of solution* is negative, i.e., heat is absorbed when the salt is dissolved. This depends upon the fact that thermal energy is needed to separate the oppositely charged ions. However, at the same time "hydration" takes place with the evolution of heat, and if this heat evolution is large enough the total heat of solution can be positive.

The *specific heat* of salt solutions is often conspicuously low, so that in many cases the salt appears to have a negative heat capacity in solution. This is due to the same factors as those which cause small or negative molar volumes in salt solutions.

(b) **The Theoretical Calculation of Activity Coefficients.** It has already been mentioned that the ions in a salt solution are not distributed in a completely random manner, but tend to arrange themselves in a certain way on account of their electric charges. A

closer study of this problem has led to a detailed theory of dilute solutions of strong electrolytes (Debye-Hückel). By applying the Boltzmann distribution law to the distribution of ions round a particular ion chosen for consideration, and then introducing Poisson's equation from electrostatics, it has proved possible to derive the following general equation for the activity coefficient,

$$-\log f = \frac{z^2 A \sqrt{\mu}}{(DT)^{\frac{1}{2}}}, \quad (53)$$

where z is the valency of the ion, μ the "ionic strength" as already defined in (19), D the dielectric constant of the medium and A a constant having the value

$$A = 1.82 \times 10^6 \quad (54)$$

if the logarithm in (53) is decadic.

For aqueous solutions and $T = 293$, $D = 80$, (52) gives

$$-\log f = 0.50 \cdot z^2 \sqrt{\mu}, \quad (55)$$

and for univalent ions

$$-\log f = 0.50 \sqrt{n_{\pm}}. \quad (56)$$

These equations agree well with the experimental values for activity coefficients derived from measurements of solubilities, freezing points and other sources.

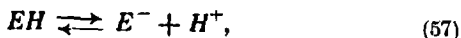
Equation (53) is however only valid at high dilutions, and in more concentrated solutions it must be modified by taking into account other factors, *e.g.*, the size of the ions.

3. WEAK ELECTROLYTES

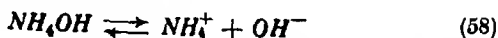
(a) **The Classical Theory.** The incomplete dissociation of salts mentioned in [2. a.] belongs to the class of phenomena which may be termed "ionic association," in which ions of opposite charge are held together by polar links (*i.e.*, electrostatic forces) to form "electrostatic molecules." While processes of this kind are predominant in solvents of low dielectric constant, in aqueous solutions we know a number of cases in which the ions are bound together by "homopolar" forces to form chemical molecules in the ordinary sense of the word. "Weak electrolytes" are compounds of this kind. The term "complex compounds" is often applied to compounds held together by homopolar forces, particularly when they may be formally regarded as salts.

The most important class of weak electrolytes comprises the

"weak acids and bases." Thus according to Arrhenius' dissociation theory a substance like acetic acid in aqueous solution is partly dissociated into acetate ions and hydrogen ions according to the equation



where E is the acetate ion. Under ordinary conditions, however, this dissociation is very incomplete, *e.g.*, in a 0.1 molar solution of acetic acid it amounts only to something over 1%. According to the same theory the weak base ammonium hydroxide dissociates similarly according to the equation



The hydrogen and hydroxyl ions are thus regarded as typical of the dissociation of acids and of bases respectively.

Investigations of the dissociation of weak acids and bases have been carried out on a large scale and have been of great importance in the development of the theory of electrolytic dissociation. It is clear that "osmotic" methods (*e.g.*, the determination of freezing points) are not suited to investigate the dissociation of these compounds, since in the case of a binary electrolyte these methods measure the quantity $1 + \alpha$ and therefore only give a very inaccurate value for the degree of dissociation α if it is small compared with unity. On the other hand, the measurement of conductivities is well suited to the problem in hand, and it is this method which has been of fundamental importance in the development of the classical theory of weak acids and bases.

Experiment shows that the conductivity of solutions of acetic acid and other weak acids shows it to be much smaller than the conductivity of strong electrolytes at the same concentration, and also to increase more rapidly with increasing dilution. Both these features agree with the assumption that weak electrolytes are only dissociated to a small extent. If the observed values are used to calculate the degree of dissociation according to the equation previously given

$$\alpha = \frac{A}{A_\infty} \quad (59)$$

and the values of α thus obtained are inserted in the equation

$$c \frac{\alpha^2}{1 - \alpha} = K, \quad (60)$$

then it is found that the last equation is satisfied with a reasonable

degree of accuracy. The simple law of mass action is thus applicable to weak electrolytes, and such electrolytes can be assigned a "dissociation constant" K which is at least approximately independent of the concentration, just as in the dissociation of gases or non-electrolytes. This will however only be true if the electrolyte is sufficiently weak.

The value of the dissociation constant K may be regarded as a measure of the strength of the acid or the base. Ostwald has measured the dissociation constants of a very large number of organic acids which were found to obey equation (60) more or less exactly. When applied to weak electrolytes, this equation is often termed "Ostward's dilution law." The application of the conductivity method will be dealt with in more detail in the chapter on electrochemistry.

(b) **The Acid-base Function. Protolysis.** We have seen that in the classical dissociation theory acids and bases were characterised by their ability to split off hydrogen ions and hydroxyl ions respectively. This characteristic may in fact be regarded as the classical definition of acids and bases. This definition does not however show any reason for the special position usually assigned to acids and bases, nor for their complementary but opposite properties. This conception also fails to account for the behaviour of acids and bases in non-aqueous media. It would appear strange in principle that properties so general as those possessed by bases should be conditioned by or related to a particular grouping of atoms like OH^- .

A more generally satisfactory description of the acid-base function is provided by the scheme

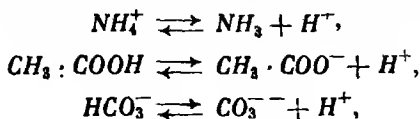


According to this definition an acid is a substance which can *lose* a hydrogen ion, thus forming a base, while conversely a base is a substance which can *gain* a proton, thus forming an acid. Two substances related by a scheme like the above are termed a *corresponding* acid and base.

The hydrogen ion H^+ is identical with the nucleus of the hydrogen atom, termed the *proton*, and since it is the simplest quantum of positive electricity which takes part in chemical reactions we shall often represent it by the symbol \oplus . Further, since the acid-base function consists in the loss and gain of such protons, it is natural to use the name *protolytes* as a term embracing both acids and bases.

It follows from (61) that corresponding acids and bases cannot

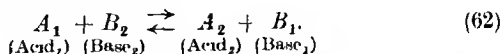
have the same charge, A always having one more positive charge than B . As examples of the acid-base function we may quote



showing that the ammonium ion is the acid corresponding to the base ammonia, while the acetate ion is the base corresponding to the acid acetic acid. According to this view the electrical charge on the molecule is thus not of primary importance for the acid-base function. According to their charge acids are termed cation acids, anion acids or neutral acids, and a similar terminology can be applied to the bases.

Both the above examples and the general scheme (61), however, represent reactions which are possible in principle rather than in practice. This is because the proton is attracted extremely powerfully by other molecules, and under ordinary circumstances cannot exist in a free state. It can therefore only be detached from one molecule if it is simultaneously transferred to a second molecule. The scheme (61) therefore does not represent processes which actually take place, but expresses the acid-base function without reference to the subsequent fate of the proton: this latter does not affect the principle involved in the definition.

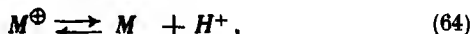
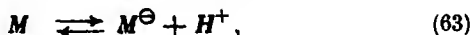
On the other hand, when considering acid-base reactions actually taking place in practice, it is of great importance to realise that the proton cannot exist in a free state, but can only be transferred from one molecule to another. Since the molecule from which it is detached is by definition an acid and the molecule to which it is transferred is by definition a base, an acid-base reaction in practice can be represented as a double reaction between two pairs of corresponding acids and bases, *i.e.*,



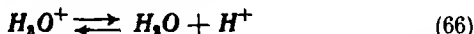
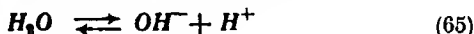
where the molecules with the same subscript represent a corresponding acid-base pair. The process in (62) is termed a *protolysis* or a *protolytic reaction*. This scheme is of very general application, since all the acid-base reactions which were treated under different headings in the classical theory—"electrolytic dissociation" of acids and bases, "hydrolysis" of salts, "neutralisation," buffer and indicator reactions, etc.—can be considered as special cases of the double acid-base reaction.

In agreement with this idea it will be seen that "electrolytic

dissociation " of acids and bases will only take place in solvents which themselves possess protolytic properties. If the molecule of the solvent is M and it has both an acid and a basic character, then its protolytic functions will be represented by the following schemes, analogous to (61),



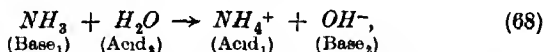
the addition of \oplus or \ominus to the molecule M representing the gain or loss of a proton respectively. Thus, for example, in the case of water, the acid and basic functions are represented by



When an acid or a base is dissolved in water the function represented by (66) or (65) respectively will come into play. For example, when acetic acid and ammonia are dissolved in water, the electrical conductivity produced must be ascribed to the two following protolytic reactions,



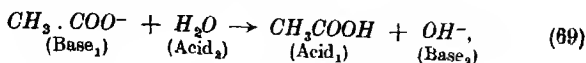
and



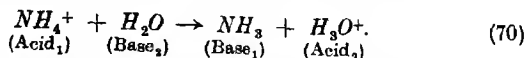
A comparison of (67) with the corresponding classical equation for the dissociation shows that the species formerly believed to be the "hydrogen ion" (and still usually referred to by this name) is actually the ion H_3O^{+} , the *hydronium* or *hydroxonium ion*, and (67) and (68) show that when acids and bases are dissolved in water, the ions H_3O^{+} and OH^{-} respectively are always formed. It is however clear that these ions are in principle on the same footing as the other acid and base molecules taking part, the special position which they may be said to occupy being entirely of a practical nature, depending on their relation to the solvent as its corresponding acid and base respectively.

If the acid or base to be dissolved in water possesses an electric charge, in which case it can only be added in the form of a salt, the ensuing protolytic reaction can be represented by equations exactly analogous to (67) and (68). In this case the protolytic reaction is often described as a "hydrolysis" of the salt in question, thus expressing the fact that the solvent plays a part in the reaction.

When sodium acetate or ammonium chloride is dissolved in water, the protolytic equations are respectively



and

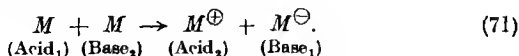


It will however appear from what has already been said that there is no difference in principle between the protolytic reactions of electrically charged and electrically neutral molecules, and it is thus superfluous to make any distinction between the two cases. Similarly, there is no reason for distinguishing between the "dissociation constant" for electrically neutral protolytes and the so-called "hydrolysis constant" for charged protolytes.

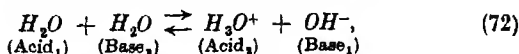
If equation (69) is read from right to left it obviously represents the reaction between acetic acid and a strong base like sodium hydroxide which is completely dissociated into ions: similarly (70) read from right to left represents the reaction between ammonia and a completely protolysed acid like hydrochloric acid. Processes of this kind are described as "neutralisation processes" and, like the reverse "dissociation" or "hydrolytic" processes, they are included in the general scheme (62). They occupy no special place among protolytic reactions, but it is convenient to use the name "neutralisation" when the protolytic reaction is almost complete, since in that case equivalent amounts of acid and base give solutions with a protolytic state near to the conventional neutral point.

Many pure liquids conduct an electric current to a small extent, and must therefore be assumed to contain ions. Apart from fused salts (which are strong electrolytes) and certain cases of electron transfer [IX. 2. j.], the formation of ions in pure liquids is due to protolytic reactions, and hence depends on the acidic and basic properties of the liquid. Protolysis will only be possible if the solvent molecules possess acidic and basic character simultaneously.

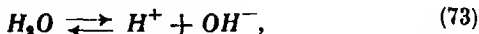
It is clear from (62) that a pure solvent which has both acidic and basic properties will undergo a partial protolytic ionisation—autoprotolysis—according to the scheme



For example, in pure water we have the equilibrium



which is generally referred to as the electrolytic dissociation of water, and is often written in the simple form



without indicating either the hydration of the proton or the basic nature of the water molecule.

If the law of mass action is applied to the above protolytic reactions of acids and bases in solution, it leads to the expression (60) previously obtained, on the basis of classical conceptions. This is most easily seen by using the fundamental formula VI. (9) and introducing the fact that the chemical potential of the solvent is constant. The importance of these laws for measuring the strength of acids and bases will be dealt with in paragraph (d).

In the case of equation (72) we obtain

$$C_{H_2O} + C_{OH^-} = K_{(H_2O)}, \quad (74)$$

The quantity $K_{(H_2O)}$ is called the dissociation constant of water, and its value has been found by various methods to be

$$K_{(H_2O)} = 0.72 \cdot 10^{-14}, \quad (75)$$

at 20°. This constant increases by about 7% for every degree rise of temperature.

These equations can always be used to calculate the hydroxyl ion concentration when the hydrogen ion concentration is given, and *vice versa*. The calculation is however only exact for solutions of very low salt concentrations, and in the presence of higher concentrations of ions it must be corrected by inserting activity coefficients.

(c) **Acidity and Basicity.** The presence of acids and bases in a solution exerts a considerable effect on many of the properties of the solution. As regards the acid-base equilibria taking part (and so some extent also in other respects) the *protolytic state* of the solution can be characterised by a single quantity, expressing the *acidity* of the solution. The definition of the acid-base function by means of (61) makes it reasonable to define the acidity by means of the chemical potential of the proton or some other quantity which varies parallel with this potential. We shall confine ourselves to considering the protolytic state at constant temperature, and will find it convenient to define the acidity as the *proton activity*. By the basicity we shall understand the reciprocal value of the acidity thus defined. The significance of these definitions is of course governed by the general considerations about ionic activities advanced in [2. b.].

In a system containing a simple protolyte system $A-B$, the

application of the thermodynamic law of mass action to the scheme (61) shows that the acidity is given by the equation

$$a_{\oplus} = K_{[A]} \frac{a_A}{a_B}, \quad (76)$$

where a is the activity and $K_{[A]}$ a constant characteristic of the protolytic system chosen, which is termed the activity constant of the acid in question.

Further, since the proton does not exist in a free state, it is seen that the presence of such a protolytic system (or *buffer system*, as it may also be called) is *necessary* to fix the acidity. The acidity of a pure acid or base, or a solution containing only an acid or a base, is indeterminate.

(76) can also be written

$$a_{\oplus} = K_A \frac{c_A}{c_B}, \quad (77)$$

where

$$K_A = K_{[A]} \frac{f_A}{f_B} \quad (78)$$

is the so-called *acidity constant*. K_A is however only constant in very dilute solutions in a constant medium, since at least one of the species A and B is always an ion and the ratio f_A/f_B therefore varies considerably with the concentration.

If for the moment we neglect such "salt effects" on the value of K_A , it will be seen that the acidity a_{\oplus} is determined by the ratio of c_A to c_B in a given protolytic system. Since K_A is at the moment undetermined, (76)–(78) can of course only be used for determining relative acidities, and for this purpose it is a matter of indifference which of the protolytic systems present is used as a basis for the determination. If the medium itself has protolytic properties (e.g., water), it is natural to choose it as the system for determining the acidity. In the case of water, (77) combined with (65) and (66) gives

$$a_{\oplus} = K_{H_2O} \frac{c_{H_2O}}{c_{OH^-}}, \quad (79)$$

or

$$a_{\oplus} = K_{H_3O^+} \frac{c_{H_3O^+}}{c_{H_2O}}. \quad (80)$$

In the absence of salt effects the acidity in a dilute aqueous solution is proportional to the hydrogen ion concentration $c_{H_3O^+}$, since in dilute solutions c_{H_2O} is constant. As already mentioned, the

absolute values of a_{\oplus} and K_{H_2O+} are unknown, and it is therefore permissible on the basis of (80) to use the convention

$$a_{\oplus} = c_{H_2O+}, \quad (81)$$

which means that the acidity of a dilute aqueous solution can be measured by the "hydrogen ion concentration" of the solution. The method of derivation used above shows that the special position thus attributed to the ion H_3O^+ in measuring the acidity of a solution is one of practical convenience rather than principle, depending on the fact that it is the acid corresponding to the solvent.

The acidity scale in other protolytically active media can be defined in an analogous way. It should however be noted that the acidities thus defined in different media are not comparable with one another.

If the solution contains a number of buffer systems, (76) will clearly be valid for each of them, so that we can write

$$a_{\oplus} = K_{[A_1]} \frac{a_{A_1}}{a_{B_1}} = K_{[A_2]} \frac{a_{A_2}}{a_{B_2}} = \dots \quad (82)$$

It was mentioned above that the product of the concentrations of hydrogen and hydroxyl ions in a dilute aqueous solution at 20° is about 10^{-14} . A solution which is 1.0 normal with respect to sodium hydroxide will therefore have a hydrogen ion concentration of about 10^{-14} , while in 1.0 normal hydrochloric acid the corresponding concentration will be 1. The hydrogen ion concentrations of solutions in common use will hence in general lie between 1 and 10^{-14} .

Aqueous solutions in which the concentrations of hydrogen and hydroxyl ions are equal are commonly termed "neutral." The concentration of each of these ions is then about 10^{-7} . If $c_{H_2O} > c_{OH}$, the solution is said to be acid or to have an acid reaction, while if the reverse is the case it is said to be alkaline or to have an alkaline reaction. All these concepts are of purely conventional origin.

On account of the large range through which the acidity can vary, it is often convenient to express it in a logarithmic scale. In this case the acidity can be measured by $\log a_{\oplus}$, termed the *acidity exponent*. If "salt effects" are neglected, we can compare the acidities of e.g. 0.1 *HCl* and 0.1 *NaOH* by saying that the acidity exponent of the first solution is 12 units greater than that of the second solution. The acidity is most commonly expressed in a logarithmic scale by specifying a quantity p_H , defined by

$$p_H = -\log c_{H_2O+}, \quad (83)$$

(Sørensen), or if necessary by the corresponding equation with a_{H_2O+}

in place of c_{H^+} . p_H is thus equal to the acidity exponent with the sign changed. Its value extends from about 0 (in 1 normal HCl) to about 14 (in 1 normal $NaOH$). In a "neutral" solution at 20°, p_H : 7.07, while "acid" solutions correspond to smaller values of p_H and "alkaline" solutions to greater values. The name "hydrogen ion exponent" is sometimes used for p_H , but since p_H is a measure of the basicity and not the acidity of the solution, this usage is somewhat inappropriate.

The protolytic state of a solution can also be given in terms of the *acidity potential*, $RT \log a_{H^+}$, which is a direct measure of the chemical potential of the proton and bears an important relation to certain electrochemical functions [IX. 2. j.].

A logarithmic scale such as the p_H scale is particularly suited to the graphical representation of variations extending over large concentration ranges. Thus if the acidity ranges from 1 to 10^{-14} in the hydrogen ion concentration scale, it will range from 1 to 14 in the corresponding p_H scale. The centre of this scale, where the concentrations of hydrogen and hydroxyl ions are equal, is termed the "neutral point." This point has however only a purely conventional significance, since the concept of neutrality is not based on any rational definition.

If the medium is protolytically inactive (e.g., carbon tetrachloride or benzene) it will be equally justifiable to choose any simple protolytic system whatever for determining the acidity on the basis of equations (76) and (77). Once this system $A \rightleftharpoons B$ is chosen, the acidity will be directly proportional to the ratio c_A/c_B . It must however be remembered that large salt effects are often present in media of this kind, so that this simple relation may only be valid in extremely dilute solutions.

The above considerations, in particular equations (76) and (77), show that the protolytic state may be fixed in practice by using a simple protolytic system $A \rightleftharpoons B$, i.e., a mixture of an acid with its corresponding base. A solution containing such a system is known as a *buffer mixture*. It is easily seen that the acidity of such a solution is not easily changed considerably, since according to (77) considerable changes in the acidity can only be brought about by a considerable alteration in c_A or c_B , which cannot be effected by adding quantities of substance which are small compared with these concentrations. The greater the concentrations of the two buffer components, A and B , the greater will be the *buffer effect*, i.e., the smaller will be the effect of added acid or base upon the acidity. However, buffer concentrations are usually kept below 0.1 molar in order to avoid changing the medium and introducing excessive "salt effects."

In representing graphically the relation between the acidity and the concentration ratio of the buffer it is convenient to replace the variables in equation (77) by p_H (or $\log a_{\oplus}$) and the acid fraction,

defined as
$$x_A = \frac{c_A}{c_A + c_B},$$

or the base fraction, defined by

$$x_B = \frac{c_B}{c_A + c_B}.$$

The relation between $\log a_{\oplus}$ and x_A or x_B is shown in Fig. 7. As the equation shows, the shape of the curves obtained is independent of the nature of the buffer system, but the position of the curve relative to a fixed point on the ordinate axis is determined by the buffer constant K_A , $\log K_A$ being the ordinate corresponding to the mid-point of the curve. The distance between the horizontal lines in the figure is one logarithmic unit. Since the buffer effect is greatest at the mid-point, the buffer mixture will be most effective in fixing the acidity of the solution when $x_A \sim x_B \sim 0.5$, i.e., when the concentrations of acid and base are approximately equal. In general, however, a buffer mixture can be used over a range of acidity of about 2 powers of ten, i.e., from $x_A = 0.1$ to $x_A = 0.9$.

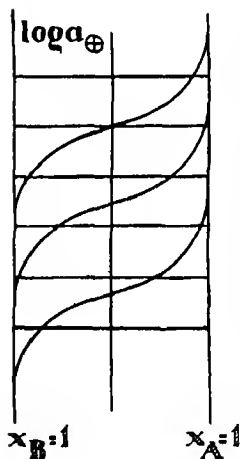


FIG. 7.

The acidity of a solution is thus most effectively fixed by using a protolytic system with a value of K_A in equation (77) which coincides with the mid-point of the acidity interval which it is desired to cover.

(d) **The Strength of Acids and Bases.** It is consonant with both common sense and tradition to measure the strength of an acid in a given medium by the acidity produced in the medium when the acid is present under given conditions, i.e., at given concentrations of the acid and the corresponding base.

The acidity constant K_A defined in (77) has this property, since

$$K_A = \frac{c_B}{c_A} a_{\oplus}. \quad (84)$$

This constant would therefore be a suitable measure of the acid

strength. Analogously, the strength of a base could be measured by the basicity constant,

$$K_B = \frac{c_A}{c_B} \frac{1}{a_{\oplus}}. \quad (85)$$

The strength constants thus defined are independent of c_A and c_B for dilute solutions in a given medium, and are obviously related by the equation

$$K_A K_B = 1. \quad (86)$$

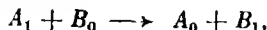
The quantity a_{\oplus} is indeterminate in the sense that its values in two different media cannot be related to one another by ordinary thermodynamic methods. In practice, therefore, we have only to deal with the relative strengths of acids and bases in the same medium. If two acid-base systems A_1-B_1 and A_0-B_0 are present in the same solution, (84) leads to the relation

$$\frac{K_{A_1}}{K_{A_0}} = \frac{c_{B_1} c_{A_0}}{c_{A_1} c_{B_0}}, \quad (87)$$

If A_0 is taken as a standard acid and its strength constant put equal to unity, the right-hand side of (87) constitutes a measure of the strength of the acid A_0 . Introducing (86), (87) can be re-written as

$$K_{A_1} K_{B_0} = \frac{c_{B_1} c_{A_0}}{c_{A_1} c_{B_0}}. \quad (88)$$

If the four concentrations entering into the protolytic equilibrium are known, (87) will give directly the relative acid strengths and (88) the product of the strength constants of the acid and base reacting. This product is identical with the mass law constant for the double acid-base equilibrium



so that we can write

$$K_{A_1} K_{B_0} = K_{A_1, B_0}, \quad (89)$$

where K_{A_1, B_0} is the mass action constant for the protolytic reaction between the acid A_1 and the base B_0 . This constant is termed the *protolytic reaction constant*, or the *protolysis constant*, and it is a quantity directly accessible to experiment, independent of any conventions. Equation (89) expresses the fact that it is equal to the product of the acidity and basicity constants for the acid and base reacting. In order to avoid the ambiguity which sometimes

arises in using this constant, we shall always let the first symbol in the subscript represent the acid, so that $K_{X \cdot Y}$ represents the protolysis constant for the acid X and the base Y .

It follows directly from (86) and (89) that

$$K_{A_1 \cdot B_0} K_{A_0 \cdot B_1} = 1, \quad (90)$$

which is a special case of the general relation

$$K_{A_1 \cdot B_2} K_{A_2 \cdot B_1} \cdots K_{A_n \cdot B_n} = 1, \quad (91)$$

which is valid for any system of n acids and their n corresponding bases. The equation also holds if some of the A and B molecules are identical.

If the medium itself is protolytically active, it is the general practice to use it as the standard acid or base in determining practical protolytic strength constants. In other words, we can use the reaction



for determining the acid strength defined as

$$K_{A \cdot M} = \frac{c_B c_{M^{\oplus}}}{c_A c_M}, \quad (93)$$

or, since c_M is constant in dilute solutions and can therefore be omitted,

$$K_{(A)} = \frac{c_B}{c_A} c_{M^{\oplus}}. \quad (94)$$

The constant $K_{(A)}$ defined in this way is identical with the *acid dissociation constant* previously introduced, and the above derivation shows that it is a suitable measure of acid strength in a constant medium which is protolytically active. If the solvent is water the constant is given by

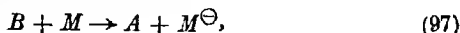
$$K_{(A)} = \frac{c_B}{c_A} c_{H_3O^+} \quad (95)$$

and if the protolysing acid is acetic acid we have

$$K_{(EH)} = \frac{c_{E^-}}{c_{EH}} \cdot c_{H_3O^+}, \quad (96)$$

where E^- is the acetate ion. This equation is identical with the classical dissociation expression (60).

The practical strength constant of a base can be determined on the basis of the reaction analogous to (92),



giving by analogy with (93)

$$K_{M \cdot B} = \frac{c_A c_{M^{\ominus}}}{c_B c_M}, \quad (98)$$

or

$$K_{(B)} = \frac{c_A}{c_B} c_{M^{\ominus}}. \quad (99)$$

identical with the classical *basic dissociation constant*. If the solvent is water the constant is given by

$$K_{(B)} = \frac{c_A}{c_B} c_{OH^-}, \quad (100)$$

and if the protolysing base is ammonia, we have

$$K_{(NH_3)} = \frac{c_{NH_4^+}}{c_{NH_3}} c_{OH^-}, \quad (101)$$

also identical with (60).

Equation (71) gives the protolytic scheme for the pure solvent analogous with (92) and (97). Analogy with (93) and (98) gives for the *autoprotolysis constant* of the solvent,

$$K_{M \cdot M} = \frac{c_{M^{\oplus}} c_{M^{\ominus}}}{c_M^2}, \quad (102)$$

or

$$K_{(M)} = c_{M^{\oplus}} c_{M^{\ominus}}. \quad (103)$$

If the solvent is water, we have

$$K_{(H_2O)} = c_{H_3O^+} c_{OH^-}, \quad (104)$$

in agreement with (74).

It will be seen from (91) or from (93) and (98) that

$$K_{A \cdot M} K_{M \cdot B} = K_{M \cdot M}. \quad (105)$$

The product of the protolytic constants for an acid and its corresponding base is thus equal to the protolytic constant of the solvent. It is easily seen by combining (94), (99) and (103) that an analogous equation holds for the dissociation constants of corresponding acids and bases, i.e.,

$$K_{(A)} K_{(B)} = K_{(M)}. \quad (106)$$

These relations (or the more general relation (91)) can be used to determine one of the constants if the values of the remaining constants are known. For example, the dissociation constant of a solvent can be calculated from the dissociation constants of acetic acid and the acetate ion dissolved in that solvent.

For a corresponding acid-base pair, the product of the acidity and basicity constants defined by (86) or of the protolysis constants defined by (90) is equal to unity. On the other hand, the strength constants defined above give a product equal to $K_{(M \cdot M)}$ or $K_{(M)}$ for a corresponding acid and base. This is because in defining the strength constants of acids and bases (equations (92) and (97)) we have used different standard systems in the two cases, namely $M^{\oplus} - M$ and $M - M^{\ominus}$ respectively. In considering values of strength constants of acids and bases it is important to note which procedure they are based upon. It may also be noted that equation (105) must be in agreement with the general equation (91), and is actually most simply derived from this general expression.

TABLE IV. *Acid Dissociation Constants.* $t = ca. 20^{\circ} C.$

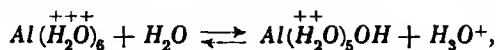
	$K_{(A)}$	$-\log K_{(A)}$
Carbonic acid 1	3.3×10^{-7}	6.48
" 2	4.4×10^{-11}	10.36
Phosphoric acid 1	7.3×10^{-3}	2.14
" 2	7.4×10^{-8}	7.13
" 3	4.1×10^{-13}	12.39
Formic acid	2×10^{-4}	3.7
Acetic acid	1.74×10^{-5}	4.76
Monochloroacetic acid	1.5×10^{-3}	2.82
Dichloroacetic acid	3.3×10^{-2}	1.48
Trichloroacetic acid	1.3×10^{-1}	0.89
Oxalic acid 1	6.5×10^{-2}	1.19
" 2	6.1×10^{-6}	4.21
Succinic acid 1	6.5×10^{-6}	4.19
" 2	2.7×10^{-6}	5.57
Fumaric acid	9.3×10^{-4}	3.03
Maleic acid	1.0×10^{-2}	2.00
Benzoic acid	6.7×10^{-5}	4.18
<i>o</i> -Hydroxybenzoic acid	1.06×10^{-3}	2.97
<i>m</i> - "	8.7×10^{-5}	4.06
<i>p</i> - "	3.3×10^{-5}	4.48
<i>o</i> -Nitrobenzoic acid	6.3×10^{-3}	2.20
<i>m</i> - "	3.5×10^{-4}	3.45
<i>p</i> - "	4.0×10^{-4}	3.40
Phenol	1.3×10^{-10}	9.89

TABLE V. *Base Dissociation Constants. $t = ca. 20^\circ C.$*

	$K_{(B)}$	$-\log K_{(B)}$
Ammonia	1.75×10^{-5}	4.76
Methylamine	4.4×10^{-4}	3.36
Dimethylamine	5.12×10^{-4}	3.29
Trimethylamine	5.27×10^{-5}	4.28
Aniline	5×10^{-10}	9.30
<i>o</i> -Chloraniline	5×10^{-12}	11.3
Pyridine	1.7×10^{-9}	8.8
Piperidine	1.6×10^{-8}	2.8

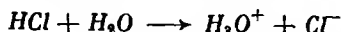
Tables IV and V give the dissociation constants as defined by equations (95) and (100) for a number of acids and bases at about $20^\circ C.$ The figures following the names refer to the stage of dissociation, *e.g.*, K for carbonic acid 2 is the dissociation constant of the HCO_3^- ion. The values of the constants will be seen to depend to a very great extent on the chemical nature of the protolyte. Certain regularities are apparent on considering the tables. Thus it will be seen that the introduction of "electronegative" atoms or radicals such as Cl , NO_2 , OH into the acid radical has the effect of increasing the acid strength and hence of diminishing the strength of the corresponding base. The distance between the proton and the group introduced also has a considerable effect. It will also be seen that the electric charge on the molecule has a large effect on the dissociation constant, an increase of negative charge increasing basic strength and decreasing the acidic strength. For example, the second stage dissociation of an acid will in general have a considerably smaller strength constant than the first stage: moreover, this charge effect is decreased if the distance between the charge and the dissociating proton is increased, as may be seen by comparing the results for oxalic and succinic acid.

The influence of the electric charge can be regarded as primarily a purely electrostatic effect, depending on the attraction between the proton and the negatively charged base. If therefore we consider a series of molecules containing hydrogen which are similar except in their electric charges, we should expect to find an increasing acidity with increasing positive charge. This is clearly shown by the acid nature of hydrated ions: *e.g.*, the increasing acidity of the hydrated sodium, magnesium and aluminium ions. In the last case the protolysis taking place in aqueous solution is represented by the equation



the dissociation constant being about 10^{-5} . The considerable acid strength exhibited by ions of this kind can be attributed to the combined effect of the high positive charge and the large number of hydrogen atoms in the ion.

Acids and bases which are practically speaking completely protolysed in water are usually termed "strong" acids and bases. Examples of strong acids are hydrogen chloride, hydrogen bromide and hydrogen iodide: thus for hydrogen chloride the protolysis



is practically complete. The actual strength may however vary very much from one "strong" acid to another, since the concentration of undissociated acid c_A which enters the equation determining the strength constant,

$$K_{(A)} = \frac{c_B}{c_A} c_{H_3O^+},$$

may vary within very wide limits even when there is practically 100% dissociation. In such cases it is usually possible to detect marked differences in the degree of dissociation of "strong" acids by using media of weaker basic character or lower dielectric constant: in this way it has been shown that the constants for the three hydrogen halides decrease in the order

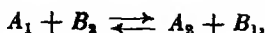
$$K_{HI} > K_{HBr} > K_{HCl}$$

No electrically neutral bases are known which are "strong" in water in the same sense as the acids mentioned above. The strongest neutral base known is piperidine, for which

$$K_{(B)} = \frac{c_{C_5NH_{11}^+} c_{OH^-}}{c_{C_5NH_{11}}} = \text{ca. } 10^{-3}.$$

On the other hand, many anion bases are known which are completely protolysed in aqueous solution.

If we have two corresponding acid-base systems in equilibrium,



then the general relation of the strength constants to the equilibrium set up is given by equation (87) or (88),

$$K_{A_1 \cdot B_2} = \frac{K_{A_1}}{K_{A_2}} = \frac{K_{(A_1)}}{K_{(A_2)}} = \frac{c_{B_1} c_{A_2}}{c_{A_1} c_{B_2}}. \quad (107)$$

If this formula is applied to the dissociation of a mixture of two acids dissolved in water, we have

$$\frac{c_{B_1}}{c_{A_1}} = \frac{x_{B_1}}{x_{A_1}} = \frac{\alpha_1}{1 - \alpha_1}, \quad \frac{c_{B_2}}{c_{A_2}} = \frac{x_{B_2}}{x_{A_2}} = \frac{\alpha_2}{1 - \alpha_2}, \quad (108)$$

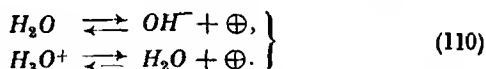
so that in this case

$$\frac{K_{A_1}}{K_{A_2}} = \frac{K_{(A_1)}}{K_{(A_2)}} = \frac{\alpha_1(1 - \alpha_2)}{\alpha_2(1 - \alpha_1)}, \quad (109)$$

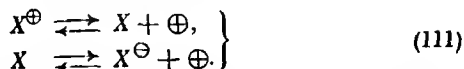
independent of the concentrations of the two acids. An analogous relation holds for a solution of two bases in water.

As in many other cases, it is often convenient to use the logarithms of the strength constants of acids and bases in place of the actual values. The logarithm of the strength constant can be suitably termed the *strength exponent*. This term is however used widely for the negative value of the logarithm, written p_K by analogy with p_H .

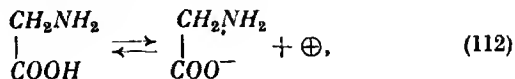
(e) **Ampholytes.** Certain substances possess at the same time both acidic and basic properties, being able both to lose and to gain a proton. Such substances are termed *ampholytes* or *amphotropic substances*. As already mentioned, water possesses this property in virtue of the following reactions,



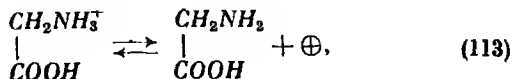
This type of behaviour can be represented generally by the symbols introduced above,



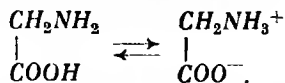
An ampholytic change can often lead to the formation of isomeric ampholyte molecules. For example, this is the case in the important group of ampholytes comprising the amino-acids. Thus amino-acetic acid can act as an acid



or as a base



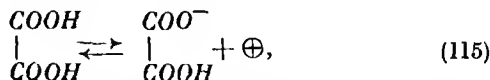
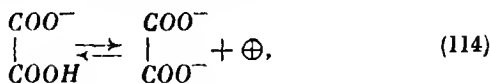
while if both these functions are exerted simultaneously, the transfer, of a proton within the molecule takes place,



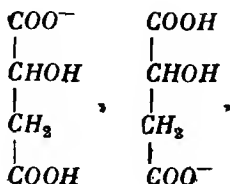
This leads to the formation of a so-called *zwitterion*, or *ampho-ion*, which is not really an ion since it has no net charge, but which has a positive and a negative charge in different parts of the molecule. The molecule thus has an electric moment, and is termed a *dipole*: the existence of such dipoles is however by no means confined to zwitterions.

According to the law of mass action, in a given medium there will be a constant ratio between the concentrations of the two forms in which the ampholyte can exist. Since in aliphatic compounds the acid strength of the —COOH group is much greater than that of the —NH_3^+ group, in the above example the ampholyte will exist predominantly as the *ampho-ion* in aqueous solution. Otherwise, however, the protolytic behaviour of ampholytes is formally independent of their special structure and possibility of isomerisation.

Just as for acids and bases in general, the net electric charge on an ampholyte molecule is not of primary importance in determining its protolytic behaviour. Thus it will be seen that the acid oxalate ion behaves as an ampholyte in virtue of the following reactions, which are quite analogous to those shown in (111),

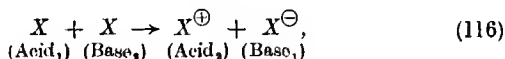


In certain cases this type of ampholyte can also occur in isomeric forms, which must be present in a fixed concentration ratio in a given medium; e.g., in the case of the acid oxalate ion we have the two forms



On the other hand, the special structure of the ampholyte molecule will of course have a great effect on the *magnitude* of the strength constants referring to its two protolytic reactions. If we assume that the acidic and basic groups which are present retain a fairly constant strength, then it is possible to deduce approximately the acidic and basic strengths of the molecule from its structure (e.g., we may calculate the constants for amino-acetic acid from the known constants of acetic acid and methylamine). Conversely, the measured constants may be used to draw conclusions about the structure of the molecule; for example, the zwitterion formula for glycine has been established in this way.

If an ampholyte X is dissolved in a protolytically inert ("aprotic") solvent, the equilibrium set up will be determined solely by the reaction



and the two species X^{\oplus} and X^{\ominus} will hence be present at equal concentrations. On the other hand, the same experiment carried out in water will in general lead to different concentrations of X^{\oplus} and X^{\ominus} , since in addition to the reaction in (116), protolysis with the solvent can also take place.

The ratio of the concentrations of X^{\oplus} and X^{\ominus} in aqueous solutions of varying acidity can be found by using the equation

$$K_{(X \cdot a)} = \frac{c_X^{\ominus}}{c_X} c_{H_2O^+}, \quad (117)$$

where the addition of a or b to the subscript X indicated that K is the acidic or basic dissociation constant of the ampholyte. We also have

$$K_{(X \oplus)} = \frac{c_X}{c_{X^{\oplus}}} c_{H_2O^+}. \quad (118)$$

Multiplying the last two equations together,

$$K_{(X \cdot a)} K_{(X \oplus)} = \frac{c_X^{\ominus}}{c_{X^{\oplus}}} c_{H_2O^+}^2. \quad (119)$$

We can now introduce (106), which in the present instance becomes

$$K_{(X \oplus)} K_{(X \cdot b)} = K_{(H_2O)} = c_{H_2O^+} c_{OH^-}, \quad (120)$$

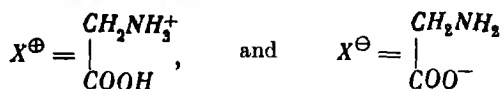
giving the following relation for determining the state of the ampholyte in a solution of given acidity

$$\frac{K_{(X \cdot a)}}{K_{(X \cdot b)}} K_{(H_2O)} = \frac{c_X^{\ominus}}{c_{X^{\oplus}}} c_{H_2O^+}^2. \quad (121)$$

There will be a certain hydrogen ion concentration at which the ampholyte is dissociated to an equal extent as an acid and as a base, so that $c_X^{\oplus} = c_X^{\ominus}$, and the total charge on the ampholyte is zero. This hydrogen ion concentration is known as the *iso-electric point* of the ampholyte, and (121) shows that it is given by

$$c_{H_2O^+} = \sqrt{\frac{K_{(X \cdot a)}}{K_{(X \cdot b)}} K_{(H_2O)}} \quad (122)$$

For example, the acidic and basic dissociation constants of glycine at 25° are respectively 1.8×10^{-10} and 2.7×10^{-12} , so that the hydrogen ion concentration at which the two ions



are present in equal concentrations is given by

$$c_{H_2O^+} = \sqrt{\frac{1.8 \cdot 10^{-10}}{2.7 \cdot 10^{-12}} 1.0 \cdot 10^{-14}},$$

or

$$c_{H_2O^+} = 8 \cdot 10^{-7},$$

In the case of electrically charged ampholytes, the state analogous to the iso-electric point will be when the total charge is equal to the charge on the unprotolysed ampholyte. The hydrogen ion concentration at this point can of course be calculated from equation (122), but in this case it is simpler to use (119) directly. Thus for the "iso-electric point" of the acid oxalate ion we find

$$c_{H_2O^+} = \sqrt{K_{(OxH_2)} K_{(OxH^-)}},$$

where the two constants are the acid dissociation constants for the first and second stage dissociations of oxalic acid. Introducing the values given in Table 4, we find

$$c_{H_2O^+} = \sqrt{6.5 \cdot 10^{-2} \cdot 6.1 \cdot 10^{-5}} = 1.99 \cdot 10^{-3},$$

as the "isoelectric point" of the ion.

(2) **Indicators.** A protolytic indicator is a simple acid-base system in which the acidic and basic form have different colours in solution. If the two forms are I_A and I_B respectively, then the scheme relating them is



which is identical with the general scheme for the acid-base function. It therefore follows that the equilibrium between indicators and other acids and bases follows the same laws as those developed above for acid-base systems in general.

The colour assumed by an indicator in solution depends partly on the total indicator concentration and partly on the ratio between the concentrations of the two forms present in solution. This ratio is determined by the equation

$$K_{IA} = \frac{c_{IB}}{c_{IA}} a_{\oplus}, \quad (124)$$

and thus depends on the acidity of the solution and the strength constant of the indicator. For a given indicator the acidity is thus determined by the indicator ratio, and can be calculated from the results of colorimetric measurements. If as before we identify a_{\oplus} with $c_{H_3O^+}$ in aqueous solution, so that K_{IA} becomes identical with the ordinary dissociation constant $K_{(IA)}$, then equation (124) can be used directly for determining hydrogen ion concentrations in aqueous solutions. $K_{(IA)}$ is obviously equal to the hydrogen ion concentration at which I_A and I_B have equal concentrations. It can be equally well expressed in a logarithmic scale.

According to (124), the effect of the acidity or the hydrogen ion concentration on the indicator ratio will be the same for all indicators. However, a change in the acidity will bring about an observable colour change only if this ratio is not too far removed from unity. In making colorimetric measurements of hydrogen ion concentrations it is therefore necessary to choose the indicator so that the value of $K_{(IA)}$ is in the neighbourhood of the hydrogen ion concentration to be measured. It may be assumed in general that useful results can be obtained if the indicator ratio is between the extremes 0.1 and 10, corresponding to a range of 2 p_H units. The specific nature of the indicator colours may however limit or extend this range.

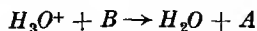
Table VI gives the names of a number of indicators together with the colours of the two forms and the p_H -range in which a change of acidity causes a marked change in the colour. In the case of two-colour indicators, the value of $K_{(IA)}$ will approximately coincide with the mean value of the hydrogen ion concentration in this range. Thus for methyl red we have approximately $\log K_{(IA)} = -5.3$, so that the acidity of a solution for which colorimetric measurements give an indicator ratio $c_{IA}/c_{IB} = 2$ is $p_H = 5$, or $c_{H_3O^+} = 10^{-5}$.

TABLE VI

Indicator.	Acid colour.	Basic colour.	p_H -range.
Methyl orange . .	red	orange	3.1— 4.4
Bromphenol blue. .	yellow	violet	3.0— 4.6
Methyl red . . .	red	yellow	4.4— 6.2
Bromcresol violet .	yellow	violet	5.2— 6.8
Bromthymol blue .	yellow	blue	6.0— 7.6
Neutral red. . . .	red	yellow	6.8— 8.0
o-Cresol red . . .	yellow	red	7.0— 8.8
m-Cresol violet . .	yellow	violet	7.6— 9.2
Thymol blue . . .	yellow	blue	8.0— 9.6
Phenolphthalein .	colourless	red	8.0— 9.8
Thymolphthalein. .	colourless	blue	9.3—10.5

(g) **Titration Theory.** In a protolytic titration the amount of acid or base is determined by the addition of a known amount of base or acid respectively until the equivalent point is reached. The titration is most commonly carried out in aqueous solution and the equivalent point determined by using a protolytic indicator. Use may also be made of the change in the conductivity or in the potential with respect to a hydrogen electrode when the solution is made to pass the neutral point by the addition of titrant solution.

Titration is based on the assumption of an almost complete reaction between the titrand and the titrant. For example, if a weak base is titrated with a weak titrant acid having the same strength as the acid corresponding to the titrand base, then at the equivalent point the protolytic reaction will only have taken place to the extent of 50%. If on the other hand the titrant acid is 10^6 times as strong as the acid corresponding to the titrand base, the extent of the protolytic reaction at the equivalent point will be 99.9%. The titrant (which can be chosen at will) should therefore always be a strong acid or a strong base, *e.g.*, hydrochloric acid or sodium hydroxide, which exist in aqueous solution as the H_3O^+ and OH^- ions respectively. Let us suppose that a weak base B is to be estimated by titration. On adding H_3O^+ , the reaction



takes place, and if an equivalent amount of titrant is added, the solution can be considered to be a solution of A in pure water, and will have the acidity of such a solution. In order to determine the point of equivalence, it is necessary to use an indicator which has a known and easily recognisable colour in a solution of this acidity. From what has been said above it will be seen that the p_H at which

the indicator changes colour must lie in the neighbourhood of the acidity of the solution of A .

For example, in titrating ammonia with hydrochloric acid it is necessary to use an indicator having an acid constant near to the acidity of the solution of ammonium ion produced. This can be calculated from the equation

$$c_{H_3O^+} = \sqrt{K_{(NH_4^+)} c_{NH_3}},$$

where $K_{(NH_4^+)}$ is calculated from the value of $K_{(NH_3)}$ in Table V by using the relation (106). If therefore the solution is 0.01 molar with respect to ammonium chloride at the conclusion of the titration, the acidity is found to be about 2×10^{-6} , so that bromocresol violet or methyl red can be used as indicator. In the same way it can be shown that phenolphthalein is a suitable indicator for titrating a weak acid like acetic acid with sodium hydroxide.

The sharpness of the colour change depends on the ratio between the relative acidity change and the corresponding quantity of titrant added. Thus if the concentration of a weak base B is being determined by means of a titrant acid A_T (which is very much stronger than the acid corresponding to B), then since the hydroxyl ion concentration at the equivalent point is infinitesimal, the following relation will hold :

$$dc_{A_T} = dc_{H_3O^+} - dc_B, \quad (125)$$

where c_{A_T} is the concentration of titrant acid which would be present if no protolytic reaction had taken place. Since

$$\frac{c_{H_3O^+} c_B}{c_A} = K_{(A)}, \quad (126)$$

and c_A is practically constant in the neighbourhood of the critical point, we have

$$dc_B = -K_{(A)} c_A \frac{dc_{H_3O^+}}{c_{H_3O^+}}, \quad (127)$$

and hence

$$\frac{dc_{A_T}}{d \ln c_{H_3O^+}} = c_{H_3O^+} + c_B. \quad (128)$$

The product of the two terms on the right-hand side of (128) is shown by (126) to be $K_A c_A$, which is practically speaking constant near the end-point: hence the sum of the two terms will be a minimum when they equal one another. The sharpest colour change, i.e., the maximum change in acidity for a given addition of titrant, will occur when $c_{H_3O^+} = c_B$, i.e., at the point of equivalence.

The above treatment refers to the case in which the reaction between the acid and base goes almost to completion, the product K_A, K_B , of the strength constants being considerably greater than unity. A very different state of affairs is encountered for acid-base mixtures in which K_A, K_B , has a value very near to unity; in this case the change of acidity caused by a given change in the ratio of acid to base has a *minimum* at the equivalent point. The special case in which A_1 and B_2 are a corresponding acid-base pair has been treated in paragraph (c).

It is easily shown that the complete expression for the accuracy of titration is given by

$$\frac{d \ln c_{H_3O^+}}{d \ln c_{A_T}} = \frac{c_{A_T}}{2c_{H_3O^+}}, \quad (120)$$

which can be used to derive the conditions for titration, assuming ideal measuring instruments.

If the base is very weak, it will not be possible to titrate it in aqueous solution. In such a case it is possible in principle to carry out the titration in a protolytically indifferent solvent instead of in water, when the extent of the protolytic reaction will depend only on the value of the product K_A, K_B , for the two protolyte systems taking part. It must however be remembered that the processes involved may depend very largely on the dielectric constant of the medium chosen.

(h) **Salt Effects.** The above treatment of weak electrolytes is based on the assumption that the solute molecules behave ideally, i.e., that they obey the laws of dilute solutions. If the ionic concentration is very low, as in solutions of weak electrically neutral acids or bases, this condition will be approximately fulfilled, which explains the applicability of the classical theory to such solutions. We have seen that in solutions of higher salt concentrations the interionic forces lead to various anomalies for strong electrolytes, and the same forces will also cause marked deviations from the simple protolytic laws in solutions of weak electrolytes in which the ionic concentrations are not very low.

The effect of salts upon equilibria between weak electrolytes can be derived from the various equations already obtained for the protolytic strength constants. Of these, (95) is identical with the classical expression for a dissociation constant. $K_{(A)}$ is therefore not a true constant, but is related to the thermodynamic constant $K_{[A]}$ by the expression

$$K_{(A)} = K_{[A]} \frac{f_A}{f_B f_{H_3O^+}} \quad (130)$$

This equation shows that the effect of salts on protolytic equilibria will depend on the charge type of the protolyte. Taking logarithms

and introducing equation (18) and the following relation derived from (61)

$$z_A = z_B + 1, \quad (131)$$

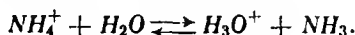
we have

$$\log K_{(A)} = \log K_{[A]} - 0.5 [z_A^* - z_B^* - 1] \sqrt{\mu}, \quad (132)$$

or

$$\log K_{(A)} = \log K_{[A]} - z_B \sqrt{\mu}. \quad (133)$$

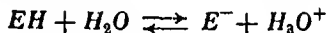
When $z_A = 1$ this gives $K_{(A)} = K_{[A]}$, showing that a dissociation or a buffer equilibrium involving an acid with a single positive charge is not displaced by the addition of salt. For example, this is true of the equilibrium



If, on the other hand, the acid is electrically neutral, *i.e.*, $z_A = 0$ and hence $z_B = -1$, we have

$$\log K_{(A)} = \log K_{[A]} + \sqrt{\mu}, \quad (134)$$

showing that in this case addition of salt causes an increase of dissociation. For example, this is the case for a solution of acetic acid, the equilibrium



being displaced to the right to an extent which can be calculated from (134). Thus if the ionic concentration in the original acetic acid solution is taken to be 0.001 and the solution is then made 0.1 molar with respect to potassium chloride, then the calculated increase in $\log K_A$ is 0.285, corresponding to an increase of about 93% in the dissociation constant of the acid. Actually, however, the observed increase is considerably less than this, since (18) is only exactly valid at ionic concentrations considerably smaller than that considered here.

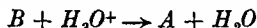
It is easily seen from the above treatment that the effect of added salt on a protolytic equilibrium will be small or zero if the two protolytic systems in equilibrium with one another are of the same charge type, while a displacement of equilibrium will take place if the charge types are different. The equilibrium will be displaced in the direction of the system which is most highly charged, though (18) shows that not only the total number of charges but also their distribution among the different species must be taken into account.

The displacement of these equilibria will of course not only be caused by the addition of a protolytically inactive salt like *KCl*, but also by alterations in the concentration of the buffer system

itself. In order to study the displacement experimentally, it is necessary to use methods which give a direct measure of the concentrations of single molecular or ionic species. Thus it may be mentioned that both the colour (light absorption) and the catalytic activity [VIII. 1. f.] are often insensitive to the addition of salts, and are thus suitable properties for determining the concentrations of the species in question.

It will also be seen that "salt effects" (and electrolyte anomalies in general) will disappear if the electrolytic equilibria are studied in a medium consisting of a salt solution with a concentration large compared with those of the reacting ions.

(i) **The Solubility of Weak Electrolytes.** A sparingly soluble base B will in general be more soluble in an acid solution than in pure water. If the base is weak and the acid is added as H_3O^+ , then the increase in solubility will be due to the reaction



since the solubility in pure water is

$$s_0 = c_B,$$

and in the acid solution

$$s = c_B + c_A = s_0 + c_A$$

We can now introduce the dissociation constant,

$$K_{(A)} = \frac{c_B}{c_A} \cdot c_{H_3O^+},$$

and the relation $C_{H_3O^+} = c_{H_3O^+} + c_A$, where $C_{H_3O^+}$ is the H_3O^+ concentration which the solution would have if no protolytic reaction took place: it is thus known from the amount of acid added. This gives

$$K_{(A)} = s_0 \left[\frac{C_{H_3O^+}}{s - s_0} - 1 \right]. \quad (135)$$

This equation makes it possible to calculate the increase in solubility $s - s_0$ when $K_{(A)}$ or $K_{(B)}$ is given. Conversely, the equation may be used to calculate the dissociation constants of weak acids and bases from the results of solubility measurements.

In general, it is often possible to use the solubility changes caused by the addition of small amounts of other substances for investigating equilibria in homogeneous solutions. It has been previously shown that the effect of one salt upon the solubility of another can often be attributed to the interionic forces [2. d.]. However, much greater effects are met with if the added salt leads to the formation of a weak electrolyte, *e.g.*, a complex salt. For example, if the solubility of cuprous chloride is measured in solutions of potassium

chloride it is found that with small *KCl* concentrations the solubility is lowered (in agreement with the principle of the solubility product), but that as the *KCl* concentration increases, the solubility first passes through a minimum and then increases rapidly. This is due to the formation of a complex ion according to the equation



Since the dissolved cuprous chloride is present partly as Cu^+ ions and partly as CuCl_2^- ions, the solubility can be written as

$$s = c_{\text{Cu}^+} + c_{\text{CuCl}_2^-}.$$

Since

$$c_{\text{Cu}^+} c_{\text{Cl}^-} = K_1,$$

and

$$\frac{c_{\text{CuCl}_2^-}}{c_{\text{Cl}^-}} = K_2, \quad (136)$$

this leads to the following equation,

$$s = \frac{K_1}{c_{\text{Cl}^-}} + K_2 c_{\text{Cl}^-}. \quad (137)$$

which can be used to determine the variation of the solubility with the concentration of *KCl*. In a quantitative treatment of equilibria of this kind it is necessary to take into account the interionic forces, i.e., to introduce ionic activities in place of concentrations in the above equations, unless the substances in equilibrium are dissolved in a concentrated salt solution in which the activity coefficients are constant.

4. THE EFFECT OF TEMPERATURE ON ELECTROLYTIC EQUILIBRIA

The effect of temperature upon the activity coefficients of ions has been dealt with in [2. g.] and [2. h.], especially equations (50) and (51). There is no real equilibrium in dilute solutions of strong electrolytes, since they are completely dissociated, and a theoretical calculation of the effect of temperature is therefore limited to heterogeneous equilibria, in particular the solubility equilibria for sparingly soluble salts.

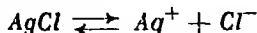
(a) **The Solubility of Sparingly Soluble Salts.** The effect of temperature on the equilibrium constant for a homo-heterogeneous process is expressed by equation VI. (44). This equation will be applicable to a solubility equilibrium if *p* is replaced by *a* and K_p by K_a . We can therefore write

$$RT^2 \frac{d \ln K_a'}{dT} = Q, \quad (138)$$

where *Q* is the heat of solution of the salt in question and K_a' is

the thermodynamic mass action constant with the term referring to the solid phase omitted.

For the solubility equilibrium



Q is the molar heat of solution of silver chloride, and

$$K'_a = a_{\text{Ag}} + a_{\text{Cl}^-}$$

the thermodynamic solubility product of silver chloride. If we neglect the variation of the activity coefficients with temperature we can write

$$\frac{d \ln K'_a}{dT} = \frac{d \ln s^2}{dT},$$

and hence

$$2RT^2 \frac{d \ln s}{dT} = Q. \quad (139)$$

Substituting in this equation the value $Q = 16,000$ cal., which is the heat evolved when one gram molecule of AgCl is precipitated from aqueous solution, we find for $t = 20^\circ \text{C}$.,

$$\frac{d \ln s}{dT} = 0.047,$$

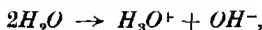
i.e., the solubility increases by about 4.6% for 1° rise of temperature.

(b) **The Dissociation of Weak Electrolytes.** In this case a treatment analogous to the above leads to the equation

$$RT^2 \frac{d \ln K_{[A]}}{dT} = Q, \quad (140)$$

where $K_{[A]}$ is the thermodynamic dissociation constant and Q is the heat of dissociation. If we again neglect the temperature dependence of the activity coefficients, $K_{[A]}$ can be replaced by $K_{(A)}$.

If (140) is applied to the protolytic reaction



putting

$$K_{(\text{H}_2\text{O})} = c_{\text{H}_3\text{O}^+} c_{\text{OH}^-} = c^2,$$

we obtain the following expression for the variation of the ionic concentration with temperature,

$$2RT^2 \frac{d \ln c}{dT} = Q, \quad (141)$$

where Q is the heat absorbed in the process represented above and c is the concentration of hydrogen or hydroxyl ions in pure water.

By measuring the heat of reaction of strong acids with strong bases we find $Q = 13,700$ cals., and insertion of this value in (141) gives at 20°C. ,

$$\frac{d \ln c}{dT} = 0.04.$$

This calculation thus shows that the ionic concentration increases by about 4% and the dissociation constant by about 8% for each degree rise in temperature.

For most weak electrolytes the heat of dissociation, and hence also the temperature variation of the dissociation constant, is considerably smaller than in the above example. Thus, for example, the dissociation of the organic carboxylic acids in general only varies slightly with temperature.

5. TERMINOLOGY OF PROTOLYTIC CONSTANTS

For convenience we give below a summary of the symbols used for protolytic constants in the above treatment.

Acidity constant

$$K_A \quad K_{A_1} \quad K_{EH} \quad K_{X \cdot a} \quad K_A = \frac{c_B}{c_A} a_{\oplus}$$

Basicity constant

$$K_{(A)} \quad K_{(A_1)} \quad K_{(EH)} \quad K_{(X \cdot a)} \quad K_{(A)} = \frac{c_B}{c_A} c_{M\oplus}$$

Acid dissociation constant

$$K_{[A]} \quad K_{[A_1]} \quad K_{[EH]} \quad K_{[X \cdot a]} \quad K_{[A]} = \frac{a_B}{a_A} a_{\oplus}$$

Base dissociation constant

$$K_B \quad K_{B_1} \quad K_{NH_2} \quad K_{X \cdot b} \quad K_B = \frac{c_A}{c_B} \frac{1}{a_{\oplus}}$$

Acid activity constant

$$K_{(B)} \quad K_{(B_1)} \quad K_{(NH_2)} \quad K_{(X \cdot b)} \quad K_{(B)} = \frac{c_A}{c_B} c_{M\oplus}$$

Base activity constant

$$K_{[B]} \quad K_{[B_1]} \quad K_{[NH_2]} \quad K_{[X \cdot b]} \quad K_{[B]} = \frac{a_A}{a_B} \frac{1}{a_{\oplus}}$$

Protolysis constant

$$K_{A_1 \cdot B_2} \quad K_{A \cdot M} \quad K_{M \cdot B} \quad K_{X \cdot M} \quad K_{M \cdot X} \quad K_{A_1 \cdot B_2} = \frac{c_{B_1} c_{A_2}}{c_{A_1} c_{B_2}}.$$

A, *B* and *X* represent respectively acid, base and ampholyte molecules. *M* is the molecule of an amphiprotic medium.

CHAPTER VIII

REACTION KINETICS

THE subject of reaction kinetics deals with the progress of chemical reactions with time. The velocity of a chemical reaction is usually taken to mean the ratio between an infinitesimal concentration change caused by the reaction and the corresponding time. For a given reaction this velocity will depend on many factors, for example the concentration of the reacting substances (when present in a homogeneous mixture), the nature of the medium in which the reaction takes place, the temperature, the catalytic effect of other substances present, etc. The objective of reaction kinetics is to determine and account for these various effects.

1. HOMOGENEOUS REACTIONS

(a) **The Order of a Reaction.** Like the laws of chemical equilibrium, the laws of reaction velocity in a homogeneous phase only assume a simple form when the substances used up in the reaction are present in a dilute state. If the velocity (as defined above) is directly proportional to the concentration of one of the substances which disappears during the reaction, so that we can write

$$-\frac{dc}{dt} = kc \quad (1)$$

where t is the time, then the reaction is said to be of the first order. If the velocity is proportional to the concentrations of two substances, or to the square of the concentration of a single substance, so that we can write

$$-\frac{dc}{dt} = kc_1c_2 \quad (2)$$

or

$$-\frac{dc}{dt} = kc^2 \quad (3)$$

the reaction is said to be of the second order, and so on for higher orders. It has been shown previously [VI. 2. b.] that we should expect a reaction of the first order when the process takes place by the spontaneous decomposition of single molecules, and a reaction of the second order when it is conditioned by a collision between

two molecules. The absolute velocity depends on the magnitude of k , which is termed the *velocity constant*. The numerical value of this constant will of course depend on the unit of time used and (except for a first order reaction) on the unit of concentration.

Equations (1) to (3) represent the velocity at an arbitrary time when the concentrations have the values inserted in the equations. To determine the order of a reaction in practice it is necessary to integrate the equations.

For equation (1) this gives

$$\ln c = -kt + i \quad (4)$$

where c is the concentration of the reacting substance left at time t , and the integration constant can be determined by applying the expression to the initial state, where $t = 0$ and $c = c_0$. This gives

$$\ln \frac{c}{c_0} = -kt \quad (5)$$

as the relation between the time and the concentration of a substance transformed by a reaction of the first order. Equation (5) can also be written

$$\ln(1 - \alpha) = -kt \quad (6)$$

where α is the extent of the reaction.

In the case of equation (3), which applies to a reaction between two substances present at equivalent concentrations, integration gives

$$\frac{1}{c} = kt + i \quad (7)$$

or, by eliminating i in the same way as above,

$$\frac{1}{c} - \frac{1}{c_0} = kt. \quad (8)$$

If the two reacting substances are present at different initial concentrations, c_1 and c_2 , then if at time t the concentration of each of them is reduced by x , (2) can be written

$$\frac{dx}{dt} = k(c_1 - x)(c_2 - x), \quad (9)$$

or

$$\frac{dx}{(c_1 - x)(c_2 - x)} = kdt.$$

This differential equation can be integrated by using the identity

$$\frac{1}{(c_1 - x)(c_2 - x)} = \left(\frac{1}{c_1 - x} - \frac{1}{c_2 - x} \right) \frac{1}{c_2 - c_1},$$

giving

$$\frac{1}{c_2 - c_1} \ln \frac{c_2 - x}{c_1 - x} = kt + 1,$$

or

$$\frac{1}{c_2 - c_1} \ln \frac{c_1(c_2 - x)}{c_2(c_1 - x)} = kt. \quad (10)$$

If $c_2 \gg c_1$, (10) reduces to the form

$$\frac{1}{c_2} \ln \frac{c_1}{c_1 - x} = kt, \quad (11)$$

which is easily seen to be identical with (5). This is in agreement with the fact that under these conditions only the smaller of the two concentrations will be appreciably altered in the reaction, which thus becomes formally of the first order.

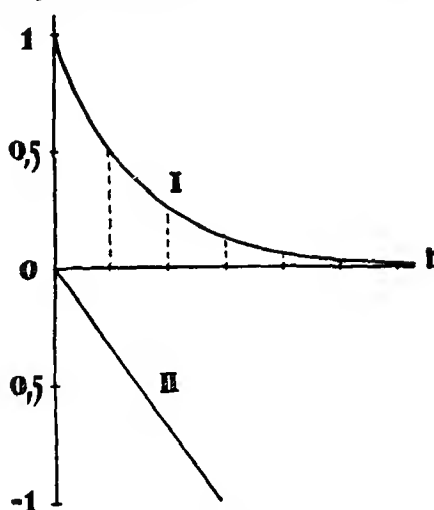


Fig 1.

There is no difficulty in obtaining expressions for processes of higher orders by an analogous procedure. Simple examples of such processes are however rare.

Special reaction mechanisms may lead to reactions with a kinetic order lower than unity [2 c.]. Such reactions are characterised by going to completion in a finite time, while for orders greater than unity the reaction is theoretically not completed until an infinite time has elapsed.

Fig. 1 shows a graphical representation of equation (5), the time t being the abscissa. Curve I shows the relation between $1 - \alpha = c/c_0$ and t , curve II the relation between $\log(1 - \alpha) = \log c/c_0$ and t . Equation (5) or (6) shows that the latter curve is a straight line. It will also be seen that a given time interval will reduce c by the same ratio, independent of the stage of the reaction. A reaction of the first order may therefore be characterised, *e.g.*, by the "half-time," *i.e.*, the time necessary to halve the concentration of the reacting substance. If (5) is written in the form

$$\log \frac{c}{c_0} = -0.4343 kt,$$

it will be seen that the half time $t_{0.5}$ is related to the velocity constant k by the equation

$$kt_{0.5} = \frac{0.3010}{0.4343} = 0.693. \quad (12)$$

The course of the reaction may therefore be equally well characterised numerically by k or by $t_{0.5}$.

(b) **Reversible Processes.** There are numerous causes which can lead to deviations from the simple behaviour described by equations (1) to (3). For example, two reactions may take place simultaneously, or the substances formed in one reaction may be transformed further into new substances (consecutive reactions). In such cases the mathematical formulation of the net reaction velocity often offers considerable difficulty. We shall consider only one possible complication, *i.e.*, the case in which a first order reaction is reversible and goes to an equilibrium state, the reverse reaction also being of the first order.

We shall assume that the two opposed reactions take place independently of one another. If the equilibrium is $A \rightleftharpoons B$ and the extent of the reaction from left to right is α , the total concentration being c_0 then the velocity h_1 of the process $A \rightarrow B$ is given by

$$h_1 = k_1 c_A = k_1 c_0 (1 - \alpha),$$

and the velocity h_2 in the reverse direction by

$$h_2 = k_2 c_B = k_2 c_0 \alpha.$$

The total velocity directly observed is thus

$$h = h_1 - h_2 = c_0 \frac{d\alpha}{dt} = c_0 [k_1 (1 - \alpha) - k_2 \alpha]. \quad (13)$$

If at equilibrium $\alpha = \alpha_0$, we have from (13)

$$k_1 (1 - \alpha_0) = k_2 \alpha_0, \quad (14)$$

since the total velocity is zero at the equilibrium point. If further we write $k_1 = k\alpha_0$ and $k_2 = k(1 - \alpha_0)$, (13) gives

$$\frac{d\alpha}{dt} = k(\alpha_0 - \alpha),$$

or

$$\ln\left(1 - \frac{\alpha}{\alpha_0}\right) = -kt = -(k_1 + k_2)t. \quad (15)$$

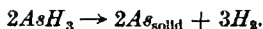
Equation (15) shows that equilibrium is reached by a reaction of the first order, the velocity constant being independent of the direction of the reaction and equal to the sum of the velocity constants of the two opposed reactions.

Equation (14) may be regarded as an expression of the chemical equilibrium in the simple system $A \rightleftharpoons B$ derived on the basis of kinetic ideas. Reference has already been made [VI. 2. b.] to kinetic laws as a general basis for the law of mass action, and it will be seen by analogy with the derivation given there that the mass law constant can be written in general as the ratio of the equilibrium constants for the two opposed processes,

$$K = \frac{k_1}{k_2} \quad (16)$$

independent of their kinetic orders.

(c) **Unimolecular Reactions.** If the process concerned consists of the spontaneous decomposition or rearrangement of a single molecule, the process is said to be *unimolecular*, and we should expect to find kinetically a reaction of the first order. As an example of a first order gas reaction following equation (5) we may mention the decomposition of arsenic hydride, which at about 300° takes place according to the stoichiometric equation



If the initial pressure is p_0 and the extent of decomposition α , then when the reaction takes place at constant volume the total pressure p corresponding to α is given by

$$p = p_0 \frac{2 + \alpha}{2} \quad (17)$$

The reaction can therefore be followed and α determined by measuring the pressure. If we introduce in (6) the values of

$$1 - \alpha = \frac{3p_0 - 2p}{p_0},$$

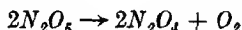
derived from (17) and the corresponding times t , a constant value is obtained for k , showing that the order of the reaction is unity.

It might therefore be supposed that the mechanism of the reaction was a spontaneous decomposition according to the scheme

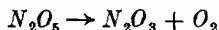


Further investigation of this reaction has however shown that the velocity constant depends on the nature of the vessel in which the decomposition takes place, and that the reaction must in reality depend upon the collision of arsenic hydride molecules with the walls of the vessel. The walls may thus be said to exert a *catalytic* effect in accelerating the reaction, which therefore does not take place by the mechanism suggested above as characteristic of unimolecular processes. On the other hand, under constant experimental conditions the catalytic effect may be regarded as constant, so that the more complicated mechanism involved in the reaction is not inconsistent with the fact that in practice the reaction appears to be of the first order when carried out under constant experimental conditions.

Other gas reactions exhibiting all the characteristics of unimolecular processes have been recently discovered. This is true, for example, of the decomposition of nitrogen pentoxide, which follows the stoichiometric equation

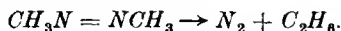


and can therefore be followed manometrically, like the last reaction. The unimolecular reaction process is probably



followed by a further rapid reaction between these primary products. The velocity of these secondary reactions has no effect on the net reaction velocity, since in a series of consecutive reactions the observed rate depends upon the rate of the slowest step.

Among other truly unimolecular processes we may mention the decomposition of acetone and ether at high temperatures, forming carbon monoxide and hydrocarbons: also the decomposition of azomethane according to the equation

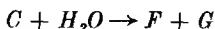


Some of these reactions, however, exhibit deviations from unimolecular behaviour at low pressures; this point will be dealt with in section (g).

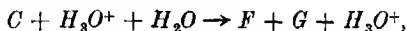
A large number of reactions in solution obey the equations for a process of the first order. For example, the decomposition of nitrogen pentoxide (mentioned above) is also a first order reaction in chloroform and similar solvents, and since the velocity is the same as it is in the gas state, the process in this case is probably

truly unimolecular. On the other hand, most of the numerous solution reactions of the first order have a higher type of mechanism but obey the first order laws in virtue of the simplifying conditions mentioned in section (a), which allow the bimolecular equation (10) to degenerate into (11).

The classical example of a first order reaction in solution is the inversion of cane sugar. The process consists in a splitting up of the cane sugar molecule into fructose + glucose, water being taken up,



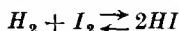
and can be followed polarimetrically, since C and $F + G$ have different optical rotations. In pure water at ordinary temperatures the velocity of the reaction is practically zero, but the addition of acids causes it to go at a measurable rate. The effect of the acids can be interpreted as *hydrogen ion catalysis*, since the nature of the chemical change is not affected by the addition, and the velocity is found to be directly proportional to the hydrogen ion concentration. Since the catalyst must be assumed to take part in the process on an equal footing with the other reacting molecules, the reaction ought presumably to be written as



so that the process is in reality a trimolecular one. However, since neither H_3O^+ nor H_2O undergo any change of concentration during the process, the course of the reaction is unimolecular, as represented by (5) or (11).

Other unimolecular reactions in solution will be mentioned in section (f).

(d) **Bimolecular Reactions.** In certain cases bimolecular gas reactions exhibit kinetic behaviour in harmony with the simple stoichiometric equation. For example, this is true both for the formation and the decomposition of hydrogen iodide,

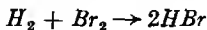


which obeys equation (10) when sufficiently far from equilibrium. The two velocity constants k_1 and k_2 are found to satisfy the relation (16)

$$\frac{k_1}{k_2} = K$$

where K is the equilibrium constant.

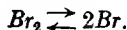
The formation of hydrogen bromide is represented by the similar stoichiometric equation



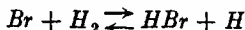
and we might therefore expect to find a similar kinetic behaviour. Actually, however, this reaction is found to obey quite a different kinetic equation, namely

$$\frac{dc_{HBr}}{dt} = \frac{kc_{H_2}\sqrt{c_{Br_2}}}{m + \frac{c_{HBr}}{c_{Br_2}}},$$

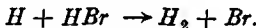
where m is a constant independent of temperature. The formation of hydrogen bromide is thus not a simple bimolecular process, and further analysis of the above result shows that it can be interpreted as a reaction between hydrogen molecules and bromine atoms, a small concentration of the latter being present in equilibrium with the bromine molecules,



The hydrogen atoms liberated by the reaction

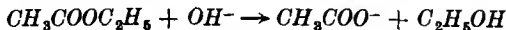


can react either with bromine molecules or with hydrogen bromide molecules, according to the equations

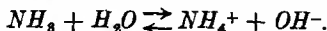


The first of these reactions accelerates the formation of hydrogen bromide, while the second retards it. All these effects are expressed in the kinetic equation given above. It is thus possible to obtain information about complex reaction mechanisms by studying the deviations from the normal kinetic equations.

Among solution reactions we may mention the saponification of an ester by a strong base like sodium hydroxide. This takes place according to the equation



and in agreement with this is kinetically bimolecular. If in place of sodium hydroxide a weak base like ammonia is used, the reaction will no longer be bimolecular, since the hydroxyl ion concentration is altered not only by being used up in the reaction, but also by the consequent displacement of the protolytic equilibrium

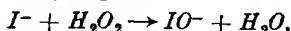


There are a large number of reactions in which the stoichiometric equation indicates a higher order, but which are actually kinetically bimolecular. This means that the reaction takes place in successive stages, and that the slow stage determining the net velocity is a

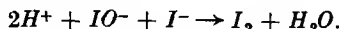
bimolecular reaction. For example, the reaction between hydrogen iodide and hydrogen peroxide can be written stoichiometrically,



which might indicate a quinquemolecular reaction. The fact that this reaction is kinetically bimolecular is explained by assuming that the rate determining step is



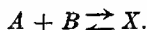
followed by the rapid reaction



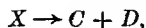
Addition of these two partial reactions gives the net stoichiometric reaction.

(e) **Ionic Reactions.** We have seen that the law of mass action for chemical equilibria can be derived from the simple kinetic equations, and the form of these equations is therefore permissible from a thermodynamic point of view when the process in question takes place in systems for which the law of mass action is valid. It has however been shown in Chapter VII, that the law of mass action is not completely applicable to electrolytic equilibria, and we should therefore anticipate that it will also be necessary to modify the kinetic equations when dealing with processes in which ions take part.

The effect of ionic charge on these equations may be derived by means of the following considerations. We will assume that two molecules A and B must form a "collision complex" X before they can react together, X being in equilibrium with A and B ,



If a small proportion of these complexes react unimolecularly to give the final products,



it is clear that the velocity of the process at a given temperature must be determined by the concentration of X .

The concentration of X can be calculated from the thermodynamic law of mass action VI. (24), giving

$$c_X = K_a c_A c_B \frac{f_A f_B}{f_X}, \quad (18)$$

where K_a is the thermodynamic mass action constant and f represents an activity coefficient. In place of the classical expression for reaction velocity,

$$h = k c_A c_B, \quad (19)$$

this treatment thus leads to the equation

$$h = k_0 c_A c_B \frac{f_A f_B}{f_X}, \quad (20)$$

or

$$k = k_0 \frac{f_A f_B}{f_X} \quad (21)$$

where k_0 is independent of the concentration.

For reactions between uncharged molecules, $f = 1$, and the classical equations are valid. The same will be true if an ion A reacts with an electrically neutral molecule B , since in this case $f_B = 1$, and $f_A = f_X$, since f is primarily determined by the magnitude of the charge. On the other hand, for a reaction between two ions the position is essentially different.

If the two reacting ions are both singly charged but of opposite sign, X will be uncharged, and hence

$$\frac{f_A f_B}{f_X} = f_1^2, \quad (22)$$

where f_1 is the activity coefficient for a univalent ion. Since this activity coefficient decreases with increasing salt concentration, the velocity constant k will also decrease, *i.e.*, there will be a *negative salt effect*. If, on the other hand, the two ions have single charges of the same sign, X will be doubly charged, and hence

$$\frac{f_A f_B}{f_X} = \frac{f_1^2}{f_2}, \quad (23)$$

where f_2 is the activity coefficient of a bivalent ion. The ratio on the right-hand side of (23) increases with increasing salt concentration, so that in this case there will be a *positive salt effect*. The magnitude of these salt effects can be obtained quantitatively for sufficiently dilute solutions by using VII. (18), giving respectively for the two cases mentioned above

$$\ln k = \ln k_0 - 2\alpha \sqrt{\mu}, \quad (24)$$

and

$$\ln k = \ln k_0 + 2\alpha \sqrt{\mu}. \quad (25)$$

Corresponding expressions are easily derived for reactions involving ions of higher charge type. The kinetic salt effects predicted by these equations are shown in Fig. 2, curves I and II referring to reaction between two singly charged ions of opposite and like sign respectively.

The salt effects described here will always be present when the salt concentration is altered, independent of whether this is done by adding salts taking no part in the reaction, or by altering the concentrations of the reacting ions themselves. It should also be noted that the prediction of zero salt effect in reactions involving at most one ion is subject to some qualification, since the alteration of the medium caused by the addition of salt is not entirely without effect from a kinetic point of view. Effects of this kind are however

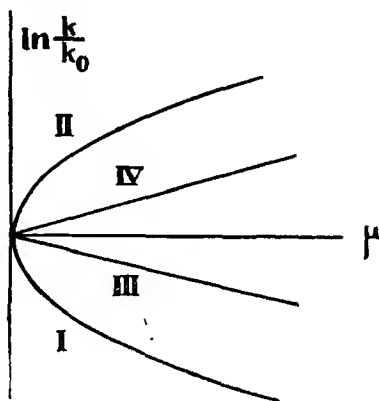


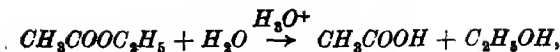
Fig. 2.

small and linear (*i.e.*, directly proportional to the salt concentration), as illustrated by curves III and IV in the figure.

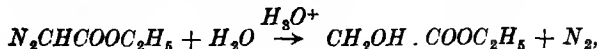
The kinetic anomalies of electrolytes can be eliminated by using as solvent a salt solution having a concentration large compared with the concentrations of the reacting ions.

(f) **Homogeneous Catalysis.** We have already mentioned the acceleration of the inversion of cane sugar by hydrogen ions in aqueous solution as an example of catalysis. This term is used in general to describe phenomena in which a reaction is accelerated by the presence of a substance (termed the *catalyst*) which takes no part in the stoichiometric equation for the total reaction. Phenomena of this kind are of frequent occurrence and considerable interest.

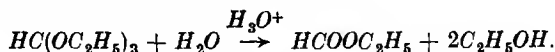
In aqueous solution the hydrogen (hydroxonium) ion and the hydroxyl ion are among the most powerful catalysts. In addition to the inversion of cane sugar, the following processes may be mentioned as being catalysed by hydrogen ions; the hydrolysis of esters,



the decomposition of diazoacetic ester,



and the hydrolysis of ethyl orthoformate



Since (apart from certain salt effects) the velocity of catalytic reactions is proportional to the concentration of the catalyst, determinations of the velocity of such reactions may be used to determine the hydrogen ion concentration of a solution. The methods used in practice for measuring the velocity depend on the nature of the reaction. The four reactions mentioned above may be most readily followed by polarimetric, titrimetric, manometric and dilatometric measurements respectively.

The hydroxyl ion is also a powerful catalyst, *e.g.*, in the hydrolysis of esters. The saponification of ethyl acetate previously described can be considered formally as an hydrolysis of the ester, catalytically accelerated by hydroxyl ions. If the velocities of hydrolysis in the presence of hydrogen and hydroxyl ions respectively are written

$$h_1 = k_1 c_{\text{Ester}} c_{H_3O^+}, \quad (26)$$

$$h_2 = k_2 c_{\text{Ester}} c_{OH^-}, \quad (27)$$

it is found that the ratio between the two catalytic constants is $\frac{k_2}{k_1} = 1,400$, so that for this reaction the OH^- ion is a much more powerful catalyst than the H_3O^+ ion.

If ethyl acetate is dissolved in pure water, in which the concentrations of OH^- and H_3O^+ ions are equal, the early stages of the hydrolysis will be predominantly due to catalysis by the hydroxyl ion. Since acetic acid is a product of the hydrolysis, the acidity will increase as the reaction progresses, causing an increase in the effect of the H_3O^+ ion and a decrease in the effect of the OH^- ion. As a result of this, the velocity constant for the hydrolysis, defined by

$$h = h_1 + h_2 = k c_{\text{Ester}},$$

will decrease, pass through a minimum, and then increase again as the reaction proceeds.

The measurement of this minimum velocity can be used to

obtain a value for the ionic product of water. The total velocity constant k can be written

$$k = k_1 c_{H_3O^+} + k_2 c_{OH^-}.$$

The product of the two terms on the right-hand side of this equation is given by

$$k_1 c_{H_3O^+} \cdot k_2 c_{OH^-} = k_1 k_2 K_{(H_2O)},$$

and is thus constant; it is then easily shown that their sum will be a minimum when the two terms are equal to one another. We therefore have at the point of minimum velocity

$$k_1 c_{H_3O^+} = k_2 c_{OH^-} = \frac{1}{2} k_m,$$

where k_m is the value of k at the minimum. This leads to

$$\frac{k_m}{2k_1} = c_{H_3O^+ (m)},$$

$$\frac{k_m}{2k_2} = c_{OH^- (m)},$$

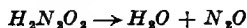
and

$$\frac{k_m^2}{4k_1 k_2} = c_{H_3O^+} c_{OH^-} = K_{(H_2O)}. \quad (28)$$

$K_{(H_2O)}$ can thus be calculated from measurements of the three kinetic constants k_1 , k_2 and k_m . The value obtained in this way agrees with the values given by other methods.

In the examples given above the catalytic effect of acids and bases in aqueous solution could be attributed to the hydroxonium and hydroxyl ions formed by protolysis. As previously explained [VII. 3], these ions do not differ in principle from other acid and base molecules, and we might anticipate that acids and bases in general would exhibit catalytic effects similar to those found for the hydrogen and hydroxyl ions. This anticipation has been realised in a number of cases.

For example, the decomposition of nitramide according to the equation



is found to be catalysed by a number of different bases. If for each base we define a catalytic constant k_B according to the equation

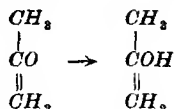
$$h = k_B c_{H_2N_2O_3} c_B$$

the value of k_B is found to increase with the basic strength K_B according to the simple relation

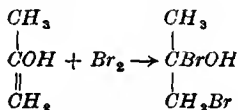
$$\log k_B = \alpha \log K_B,$$

where α is a proper fraction. Similar relationships have been found to hold in other cases, *e.g.*, the bromination of acetone and the mutarotation of glucose.

The first of these two reactions takes place when bromine is added to an aqueous solution of acetone, and is catalysed both by acids and by bases. Since the velocity is independent of the bromine concentration, bromine cannot play any part in the rate-determining process, and it must be assumed that this process consists in the enolisation of acetone according to the equation

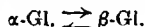


after which bromine adds rapidly on to the double bond formed,



When no bromine is present, enolisation leads rapidly to an equilibrium state in which the ketoform preponderates.

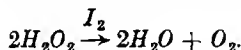
The mutarotation of glucose consists of the transformation of α -glucose to β -glucose, or the reverse reaction. This reaction is also reversible, and leads to an equilibrium state



containing about 36% of the α -form. The course of the reaction agrees with the equation (15) derived for such reversible unimolecular processes. Like the bromination of acetone, this reaction is catalysed not only by hydrogen ions and hydroxyl ions, but by both acids and bases in general.

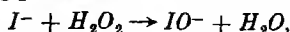
This type of catalysis, caused by acid and base molecules, is termed *protolytic catalysis*.

Among other catalytic reactions in aqueous solution we may mention the catalytic effect of iodide ions on the decomposition of hydrogen peroxide in "neutral" solutions,

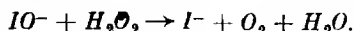


The velocity of this reaction is found to be directly proportional to the concentrations of hydrogen peroxide and iodide ion, and is numerically identical with the velocity of the above-mentioned reaction between iodide ion and hydrogen peroxide in acid solution,

leading to the formation of iodine. We must therefore assume that the rate-determining process is the same in the two cases, namely,



but that in the absence of hydrogen ions the subsequent rapid reaction is



The iodide ion is thus regenerated, and may be regarded as a catalyst since it takes no part in the net stoichiometric reaction. In contrast to most catalytic reactions, the mechanism is in this case easily understood.

Metallic ions which can have two or more valencies often exert a marked effect upon the velocity of oxidation and reduction reactions, thus exhibiting a certain analogy with protolytic catalysis. Mention may also be made of the very specific catalytic effects exerted by enzymes in biological processes, *e.g.*, the effect of pepsin in splitting up proteins.

In dealing with catalytic effects it must be specially noted that in reversible processes the two opposed reactions must be accelerated to the same extent by the addition of the catalyst, since otherwise the equilibrium would be displaced, contrary to the laws of thermodynamics. This is true both for the homogeneous catalysts dealt with here and the heterogeneous catalysts described in the next section.

(g) **Application of the Kinetic Theory of Gases.** If a reaction takes place by the collision of two molecular species *A* and *B* in a gas mixture and is found to occur with a measurable velocity, then it can be shown that only a very small proportion of the collisions actually taking place between *A* and *B* molecules result in reaction. The variation in the effectivity of different collisions is easily understood in terms of Maxwell's distribution law. In order that the reaction may take place it is generally necessary to supply energy, so that only those molecules which happen to possess a particularly high energy will be able to react. The greater the energy necessary, the more rarely will a molecule possess it, and hence the slower will the process take place in general. It is also clear that the removal of collision complexes of high energy may bring about deviations from Maxwell's distribution law in the gas mixture.

According to this view, the velocity of a bimolecular process should be represented by the product of a "collision number" (*i.e.*, the number of collisions per second between *A* and *B* molecules), and a factor expressing the probability of reaction. According to the above considerations reaction will only take place on collision if the two molecules are "activated," *i.e.*, have acquired energy

exceeding a certain *critical energy* or *activation energy*: in this case the probability factor can be calculated as a function of this energy.

It has already been shown that the number of molecules having energy in two dimensions greater than a fixed value E_k is given by III. (40). If E_k is taken to mean the energy per gram molecule, this equation becomes

$$N = N_0 e^{-\frac{E_k}{RT}}, \quad (29)$$

where N_0 is the total number of molecules and R the gas constant. The quantity we wish to find is the probability that the energy of two colliding molecules exceeds the critical value E_k : however, it is reasonable to assume that the only energy which plays any part in the reaction is the component of kinetic energy along the line of collision, i.e., in one dimension. Since the distribution of energy in one direction, e.g., along the x -axis, is independent of the energy of molecules along the y - and z -axes, the probability that a molecule has energy E' along the x -axis and simultaneously energy E'' along the y -axis will be the same as the probability that one molecule has energy E' along the x -axis and a second molecule simultaneously energy E'' along the x -axis. The distribution of energy for two molecules in one dimension is thus the same as the distribution of energy for a single molecule in two dimensions. The above equation will therefore represent the number of collision complexes whose total energy in the direction of collision (i.e., along the line of centres) exceeds the value E_k , N_0 being the total number of collision complexes. The probability of reaction is thus $e^{-\frac{E_k}{RT}}$,

and if Z is the collision number per unit volume, the calculated reaction velocity is

$$h = Z \cdot e^{-\frac{E_k}{RT}}. \quad (30)$$

This equation can also be derived more rigidly by the following considerations. In order that reaction shall take place it is necessary that in the collision the *relative velocity* of the molecules along the line of centres shall exceed a certain fixed value. According to the kinetic theory of gases the number of collisions Z_k which satisfy this condition is related to the total number of collisions Z by the equation

$$Z_k = Z e^{-\frac{\mu w_k^2}{2RT}}, \quad (31)$$

where w_k is the critical velocity and

$$\mu = \frac{M_1 M_2}{M_1 + M_2},$$

the so-called *reduced mass*.

If we wish to introduce the energy in place of the square of the velocity w_k^2 in equation (31), it must be remembered that the sum of the energies of the two moving molecules is determined not by the relative velocities but by the absolute velocities of the two molecules. However, since only motion relative to the centre of gravity of the two molecules is of importance for the energy of collision, we can assume that this centre of gravity is at rest, and the velocities w_1 and w_2 calculated on this assumption must lead to the correct energy values. The condition that the centre of gravity is stationary is

$$M_1 w_1 + M_2 w_2 = 0,$$

while the relative velocity w_r is given by

$$w_1 - w_2 = w_r.$$

This gives for the critical energy total

$$E = E_1 + E_2 = \frac{1}{2} M_1 w_1^2 + \frac{1}{2} M_2 w_2^2 = \frac{1}{2} \mu w_r^2,$$

which on insertion in (31) leads to an expression agreeing with (30).

In order to determine the important quantity E_k , the critical energy or activation energy, we can use equation (30) directly or rewrite it in the form

$$h = k c_A c_B = Z_1 c_A c_B \sqrt{T} \cdot e^{-\frac{E_k}{RT}}, \quad (32)$$

introducing the fact that the collision number is proportional to the molecular velocities, and hence [III. (8)] to the square root of the temperature. Z_1 in (32) is thus a quantity independent of temperature but dependent on the mass and size of the reacting molecules, and is clearly equal to the collision number at unit concentration and unit temperature. By means of this equation E_k can be calculated directly from velocity measurements at different temperatures without any knowledge of Z or Z_1 : thus by taking logarithms and differentiating we obtain

$$E_k = RT^2 \frac{d \ln k}{dT} - \frac{1}{2} RT. \quad (33)$$

Having thus obtained the values of the activation energy, an approximate test of equation (30) can be carried out by neglecting the variation of the collision number with temperature, since the effect of this term is very small compared with the effect of the probability factor in determining the total variation of the velocity with temperature. If we also neglect the differences arising from the different masses and sizes of the molecules, Z can be taken as constant for given values of the concentrations. Considering a series of bimolecular reactions, according to (30), the temperature

at which each reaction attains an arbitrarily chosen velocity should bear a constant ratio to the value of E_k for the reaction. This prediction is fulfilled with fair accuracy for a number of well-investigated bimolecular gas reactions at temperatures between 400° and $1,000^\circ$ absolute.

This agreement shows that the probability factor

$$e^{-\frac{E_k}{RT}}$$

is the most important variable in the expression for the velocity of bimolecular gas reactions. In order to calculate the absolute velocity from equation (30), the collision number Z must be calculated from the kinetic theory of gases. In this respect also the equation shows on the whole satisfactory agreement with experiment.

Deviations will arise if the reaction requires a closely oriented collision between the molecules, since this condition will reduce the velocity below the value given by (30). There will also be deviations if other forms of energy besides translational energy are of importance in the reaction. Further, if the reaction takes place in solution instead of in the gas phase, several of the conditions assumed in the above deduction are not fulfilled. It should however be mentioned that there are a number of cases of solution reactions for which equation (30) is a good approximation.

The kinetic theory of gases has also led to important results in treating unimolecular processes. The essential characteristics of these processes may be represented by the equation

$$-\frac{dc}{c} = kdt,$$

or, in other words, we can define a unimolecular process as one in which the probability that a molecule will react in a certain time interval is not affected by changes in the concentration or other factors at constant temperature. According to this definition, radioactive disintegration is a typical unimolecular reaction, since disintegration is a spontaneous reaction taking place in the atomic nucleus and is unaffected by external conditions. However, the nuclear character of this process makes it also independent of the temperature, thus differing from chemical unimolecular reactions.

Considerations analogous to those advanced for bimolecular reactions make it reasonable to assume that the velocity of a unimolecular chemical reaction can be obtained by multiplying the concentration of the reacting species by a probability factor, it being assumed that a molecule (like the collision complex in bimole-

cular reactions) must have a certain critical activation energy before it can react. Since unimolecular reactions do not come to a standstill when the molecules which are activated at a given moment have reacted, it must be assumed that activation by collision is continually taking place. This means, however, that the activation is really due to the same causes in both unimolecular and bimolecular reactions, thus raising the question of how the collision mechanism can differ in the two cases so as to give rise to the two different types of kinetics.

This distinction can be explained by attributing very different probabilities of reaction to the collision complex in the two cases. If the probability of reaction is great, so that in almost all cases the active complex will have reacted before undergoing further collisions which may de-activate it, then the reaction will follow a bimolecular law. In this case the occurrence of the chemical reaction will cause the whole system to deviate from Maxwell's distribution law, on account of the removal of collision complexes and molecules of high energy. If on the other hand the probability of reaction is small, the activated complex will nearly always become de-activated and will only undergo reaction in exceptional cases. The distribution law will then be practically unaffected, since the number of molecules reacting in unit time is only a small fraction of the number which are normally activated and de-activated in the same time. Since the normal energy distribution is independent of the pressure, the relative number of activated molecules will also be independent of the pressure. Since the activated molecules must be assumed to have a constant probability of reaction, this will lead to the simple first order relation for the velocity.

This interpretation in terms of the kinetic theory implies that the order of a reaction is not entirely determined by its chemical nature, but also depends on the conditions under which the process takes place. Thus it is possible that a reaction which follows a unimolecular course at a certain pressure may approach the bimolecular type as the pressure is decreased. This is because the condition for a unimolecular reaction, that the probability of reaction is small compared with the probability of collision, is no longer satisfied when the pressure and hence the collision probability decreases. It is an important confirmation of the theory that several of the unimolecular gas reactions mentioned above do in fact exhibit this type of pressure dependence.

Many unimolecular gas reactions differ from bimolecular reactions in an important respect. If we calculate the activation energy from equation (33) in the same way as for bimolecular reactions, and then calculate the absolute velocity from (30), the values obtained are often

much less than the velocities observed experimentally. We may conclude from this that the energy necessary for unimolecular activation may be present in forms other than translational, which was the only form of energy taken into account in deriving equation (30). It must be assumed that rotational and vibrational energy (which have already been seen [III.] to contribute to the specific heat and energy content of a gas) can also take part in the chemical reaction of molecules, and that this effect will become more prominent with an increase in the number of degrees of freedom involved in the molecule. In agreement with this idea, unimolecular reactions only appear to take place with relatively complex molecules in which such forms of energy can readily be stored up. It is also easy to understand why the decomposition of such molecules should be unimolecular, since the concentration of energy in the part of the molecule where it can lead to decomposition may be regarded as a type of activation having a small probability when the energy must be collected up from several degrees of freedom.

(h) **The Effect of Temperature.** A rise of temperature will in general lead to a large increase in the velocity of a chemical reaction, the temperature dependence of reaction velocity being formally similar to that of vapour pressure. By measuring the temperature coefficients of numerous reactions in the neighbourhood of room temperature, it has been found that the reaction velocity is in general roughly doubled for a temperature increase of 10° . Further, accurate measurements of both bimolecular and unimolecular reactions have shown that the velocity constant is usually represented by the following simple equation,

$$\ln k = \alpha - \frac{\beta}{T}. \quad (34)$$

On account of the relation between the mass action constant of a chemical equilibrium and the velocity constants for the two opposed reactions,

$$K_c = \frac{k_1}{k_2}, \quad (35)$$

we can write

$$\frac{d \ln k_1}{dT} = \frac{E_1}{RT^2} + i, \quad \frac{d \ln k_2}{dT} = \frac{E_2}{RT^2} + i, \quad (36)$$

these two equations leading to

$$\frac{d \ln K_c}{dT} = \frac{Q}{RT^2},$$

where

$$Q = E_1 - E_2. \quad (37)$$

These thermodynamic considerations give no information about

the quantity i appearing in the equations for the forward and reverse reactions, so that thermodynamics gives no definite prediction of the effect of temperature on the velocities.

It is however easy to calculate the effect of temperature from the kinetic theory treatment of reaction velocity developed in section (g). If we neglect the variation of the collision number with temperature (which we have seen to be of small importance) the logarithmic form of (30) and (32) gives

$$\ln k = \alpha - \frac{E_k}{RT}, \quad (38)$$

where α is a constant. This is identical in form with the empirical equation (34), and $-E_k/R$ may be obtained graphically as the slope of the straight line representing the relation between $\ln k$ and $\frac{1}{T}$.

According to equations (36) and (37), the heat evolved in the reaction $A \rightarrow B$ is equal to the difference between the heats of activation for the forward and reverse reactions. We have assumed above that bimolecular reactions take place through collision complexes, and we may picture a "critical complex" representing the transition from one system to the other. This critical complex X is common to the two systems, so that the rate-determining processes are actually



and



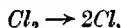
The energy of this complex must be independent of its origin, *i.e.*, whether it is formed from the A -system or from the B -system. In simple bimolecular reactions these two systems have the same kinetic energies, and the fact that different amounts of energy are necessary to convert them to the critical complex is due to the difference in the potential energies of the two systems, which in simple cases is equal to the heat evolved in the total reaction.

The reaction having the greater activation energy will have the greater temperature coefficient of reaction velocity: this kinetic result corresponds to the thermodynamic law relating the heat evolved in a reaction to the temperature dependence of the equilibrium.

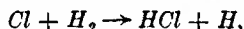
(i) **Chain Reactions.** A number of reactions exhibit the simple behaviour expected from the above mechanisms, and in other cases the total course of the reaction can be explained by assuming them to consist of several such simple reactions. There is however

a large class of reactions in the gas phase (and to a lesser extent in solution) which differ markedly in their general characteristics from those described above, *e.g.*, they are extremely erratic and difficult to reproduce, as well as very susceptible to both positive and negative catalytic effects. They also frequently exhibit at the beginning of the reaction an induction period, during which nothing appears to take place. As examples of this class of reaction we may mention the combination of hydrogen and chlorine gases, and the oxidation of sodium sulphite by oxygen in aqueous solution.

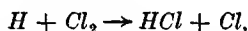
As a general explanation of this type of reaction it is assumed that molecules which have become activated by collision or in some other way (*e.g.*, by radiation, *cf.* [3. b.]) can activate new molecules when they undergo reaction; these new activated molecules are now able to react with the production of yet more activated molecules, and so on. For example, in the combination of hydrogen and chlorine the primary process is assumed to be dissociation of the chlorine molecule,



after which the active chlorine atom reacts with hydrogen,

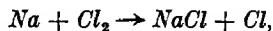


and the hydrogen atom with chlorine,



and so on. Once this chain of reactions is initiated, it will continue until all the hydrogen and chlorine have reacted, unless the active atom is de-activated by some other reaction. The net velocity of the process will clearly depend upon the number of chains and the average chain length, and the latter quantity will depend upon the probability that the chain will be interrupted by a side reaction. It is thus easy to understand how small quantities of impurities which are able to interrupt the chains can have a very great effect on the velocity of chain reactions.

The formation of hydrogen chloride in a mixture of hydrogen and chlorine is powerfully catalysed by sodium vapour, on account of the reaction



and is strongly inhibited by oxygen, which combines with the active hydrogen atoms. In mixtures free from oxygen the chain length is of the order of 10^4 molecules, and the chains are chiefly broken by the recombination of chlorine atoms, $2Cl \rightarrow Cl_2$.

In the oxidation of sodium sulphite solution, Cu^{++} ions have been found to have a very great positive catalytic effect. Small

quantities of alcohol, on the other hand, have an inhibiting effect, being oxidised to aldehyde and thus breaking the chains.

The theory of chain reactions has proved of great importance in interpreting a number of chemical phenomena, *e.g.*, explosions and combustion in flames.

2. HETEROGENEOUS REACTIONS

If molecules which can react together are present in different phases, reaction will not be able to take place until the phases in question are brought into contact. If the system still remains heterogeneous, the reaction can be considered to take place in a boundary layer, into which the reacting substances pass by diffusion. If the velocity of the chemical reaction is great, the layer in which the reaction takes place will be thin, and the velocity of the whole process will be determined by the rate at which the molecules diffuse up to the boundary layer. If on the other hand the chemical reaction is slow, the rate of diffusion is immaterial, since there will always be saturation or partition equilibrium between the phases. Under these circumstances two cases may arise. In the first case the reacting molecules are distributed through the whole system, or through a whole phase, and the velocity is determined by the ordinary homogeneous kinetic laws. In the second case the reaction is caused by direct collision between the molecules in one phase and the surface molecules of another phase; this may be regarded as the limiting case of reaction in a boundary layer, in which the thickness of the layer is reduced to molecular dimensions. Examples of all three cases are met with among heterogeneous reactions.

(a) **Diffusion Processes.** Diffusion in solution is governed by Fick's law,

$$dn = DO \cdot \frac{dc}{dx} dt, \quad (39)$$

where O is the cross-section of the diffusion cylinder in which the experiment takes place, x is the distance along the axis of the cylinder, dn the amount of substance diffusing through O in time dt , and D a constant characteristic of the substance and the solvent. This law can be applied to processes consisting in the simple solution of a solid in a liquid, it being assumed that saturation equilibrium is set up instantaneously in the liquid layer in the immediate neighbourhood of the boundary surface. If a solid is in contact with a solvent under constant stirring conditions, a stationary state will be established in which the substance passes from the surface of the solid into the homogeneous solution through a

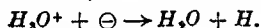
boundary layer of constant thickness but decreasing concentration. For a steady state (39) shows that the concentration gradient must be constant through the whole thickness of the boundary layer, so that the expression for the rate of solution assumes the form

$$\frac{dc}{dt} = k(s - c), \quad (40)$$

agreeing with the expression for a unimolecular process. This equation seems to be generally valid for simple solution processes.

Such reactions are characterised by the effect of the speed of stirring. Further, since the temperature coefficient of diffusion coefficients is small compared with that of ordinary chemical reactions, reactions which depend essentially on a diffusion process may be recognised by their small temperature coefficients.

(b) **The Dissolution of Metals in Acids.** Metals may be regarded as built up essentially of positive metal ions and electrons. When metals dissolve in acids, the simplest assumption is that a reaction takes place between the hydroxonium ions (or possibly other acid molecules) and the electrons in the metal,



Since the process removes a negative charge from the metal, a positive metal ion will then pass spontaneously into solution, so as to restore electrical neutrality. The primary process should belong to the third type of heterogeneous reactions mentioned above, though it will of course be followed by combination of the hydrogen atoms to give molecules.

In the case of strongly electropositive metals which dissolve rapidly in aqueous solutions of acids, the velocity of solution will undoubtedly be determined by the rate at which the acid molecules (or ions) diffuse up to the surface of the metal. In other cases the velocity may be determined by the probability of the process in the above equation. This probability depends not only on the tendency of the metal to lose an electron, as expressed by its position in the electrochemical series [IX. 2. h.], but also on the so-called "over-potential" [IX. 2. q.], which is connected with the possibility of hydrogen existing in an atomic state in the metal, and the related probability of hydrogen molecule formation.

It is partly for this reason that the process of solution is often greatly affected by the degree of purity of the metal. If electrically conducting impurities of low over-potential are present as a heterogeneous phase, the process of solution will be greatly accelerated, as we may assume that when the heterogeneous metal comes into contact with the acid solution, a short-circuited galvanic element is set up. This so-called "local element" theory explains in a satisfactory manner the behaviour of metals such as industrial iron and zinc on treatment with acids.

(c) **Heterogeneous Catalysis.** The velocity of a reaction taking place in a homogeneous phase is often greatly increased by the

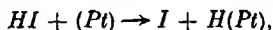
addition of certain solid substances, known as "heterogeneous" or "contact" catalysts. Well-known examples of this kind of catalytic action are the effect of platinum on the formation of water from hydrogen and oxygen, the decomposition of hydrogen peroxide in the presence of manganese dioxide, and the use of nickel in the hydrogenation of unsaturated organic compounds. For gas reactions in particular contact catalysis is often exhibited by the walls of the vessel in which the reaction takes place, so that different velocities are observed in vessels of glass, porcelain and platinum. In such cases we may speak of a "wall reaction." The decomposition of arsenic hydride has already been mentioned as an example of this type of reaction. Another example is the reaction between ethylene and bromine vapour, which takes place rapidly at ordinary temperatures in a glass vessel, but is almost completely stopped by coating the glass with paraffin wax.

The catalytic effect of solids is often highly specific, *e.g.*, a substance can be decomposed in entirely different ways by using different solid catalysts. Contact catalysis thus does not consist entirely in a change in the velocity of the reaction, but may also change the nature of the reaction taking place. These general characteristics of heterogeneous catalysis are of great importance both from a practical and from a theoretical point of view.

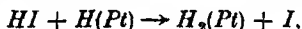
In the cases considered here the catalytic reaction is completely heterogeneous, *i.e.*, it is not necessary for the catalyst to dissolve in the surrounding medium, and the reaction must take place on the actual surface of the catalyst. It is therefore natural to explain the catalytic effect by supposing that one or more of the reacting substances are bound to the surface by the action of forces effective over very short distances, and that the deformation caused by these forces renders the molecules more able to react. It is known that substances may be bound to a solid surface in this way by the action of both chemical and physical forces; this phenomenon is known as "adsorption," and will be dealt with further in a later chapter [XI. 1]. Under certain conditions of quantity of catalyst and concentration of substance, the adsorbed substance may cover the whole of the catalyst surface, this surface being saturated with reactant. If now the adsorbed molecules react unimolecularly, the reaction velocity will not depend on the concentration of reactant, being determined only by the number of adsorbed molecules. We therefore obtain a constant velocity, independent of the extent of the reaction; this is described as a *zero order reaction*, and may be exemplified by the decomposition of hydrogen iodide on a gold surface, or the decomposition of ammonia on a heated tungsten wire. If adsorption only takes place to a small extent, the number

of adsorbed molecules may be directly proportional to the concentration in the gas or solution, and a *first order reaction* is obtained. This is the case, *e.g.*, for the decomposition of hydrogen iodide in contact with platinum. In intermediate cases the order of the reaction may be expressed by a proper fraction, corresponding to the exponent in the "adsorption isotherm" [XI. 1. a.]. It is clear that there will be a still greater number of possibilities for processes involving the reaction of molecules of different species in the adsorbed layer.

However, in order to interpret the kinetics of heterogeneous catalysis it is not always necessary to assume that the reaction takes place in the adsorbed layer. In particular, when the process follows a simple first order law it can also be interpreted as a simple collision reaction. For example, in the decomposition of hydrogen iodide as platinum, the rate-determining step is assumed to be the collision of an *HI* molecule with the platinum surface, giving rise to the reaction



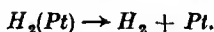
where *H(Pt)* represents atomic hydrogen combined with or dissolved in the platinum. This primary process is followed by a *second* collision, giving



after which the two liberated iodine atoms combine



and the dissolved hydrogen comes out of solution,

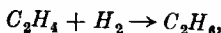


If it is assumed that the first of these four reactions is the one which determines the rate, then the above reaction mechanism will lead to a unimolecular process, in agreement with experiment. This may be contrasted with the *homogeneous* decomposition of hydrogen iodide, which is bimolecular, as previously mentioned.

The assumption of a number of successive reactions in a case like the above explains not only the order of the reaction, but also the fact that the reaction is able to take place at a relatively low temperature in spite of the high activation energy of the homogeneous reaction. In the above mechanism the reaction is split up into steps with lower activation energies, thus leading to a correspondingly higher reaction velocity (*cf.* 1. g.).

The assumption that one or more of the reactants is adsorbed has also been of great importance in interpreting heterogeneous catalysis of reactions between two or more molecules. According

to the nature of the reaction, widely different types of behaviour are met with. As an example we may mention the reaction between ethylene and hydrogen,



which is catalysed by various metals. The reaction takes place when two molecules bound to the surface of the metal in a uni-molecular layer [XI. 1. a.] come into contact. At low temperatures with copper as catalyst the velocity is given by

$$h = k \frac{c_{H_2}}{c_{C_2H_4}},$$

indicating that the metal surface is almost completely covered by ethylene molecules, so that the number of points at which a hydrogen molecule can be attached to the surface is inversely proportional to the ethylene pressure. With nickel as catalyst the adsorption behaviour is reversed, and the observed velocity expression

$$h = k \frac{c_{C_2H_4}}{c_{H_2}}$$

can be interpreted in an analogous way. At higher temperatures the adsorption is much smaller and the greater part of the metal surface uncovered, so that we find agreement with the ordinary bimolecular expression

$$h = c_{H_2} c_{C_2H_4}.$$

In other cases, *e.g.*, in the synthesis of sulphur trioxide,



with a platinum catalyst, the reaction products have a strongly inhibiting effect on the reaction. It may therefore be concluded that the reaction products are preferentially adsorbed, thus preventing fruitful encounters between the reacting molecules on the surface of the catalyst.

3. PHOTOCHEMICAL REACTIONS

Light often exerts a considerable effect upon chemical reactions. The term "light" is used here to mean electromagnetic radiation in general, but for reasons which will appear later chemical reactions are particularly sensitive to radiation of wavelength corresponding to ultra-violet or short-wave visible light. The effect of radiation upon chemical processes depends upon the adsorption of radiation energy by the system. Photochemical processes may be catalytic in the sense that the light accelerates a natural process which is

capable of taking place spontaneously. The reaction between hydrogen and chlorine is an example of this type of effect. On the other hand, it is possible for light to bring about a reaction in a direction opposite to that indicated by the affinity relations, thus indicating clearly the energetic nature of the process. An important example of this is the photochemical assimilation of carbon dioxide by plants.

(a) **Absorption of Light.** When monochromatic light passes through a homogeneous phase by which it is absorbed, it is easy to deduce from simple assumptions that the following absorption law will hold,

$$-d\ln I = kdl, \quad (41)$$

where I is the intensity of the light after passing through a layer of thickness l , and k is a constant known as the absorption or *extinction coefficient*. Integration gives

$$\ln \frac{I}{I_0} = -kl, \quad (42)$$

where I_0 is the initial intensity. The decrease in intensity is thus $I_0 - I$, and the amount of light absorbed in time t is $(I - I_0)t$. It may be noted that equation (42) is formally identical with the expression for the kinetics of a unimolecular reaction.

If the phase in question is a solution containing the absorbing substance at concentration c , (42) may be re-written as

$$\ln \frac{I}{I_0} = -klc, \quad (43)$$

where k is the *molar extinction coefficient*. This expression represents the Lambert-Beer absorption law. k is not always absolutely constant, but is usually practically speaking so unless the absorbing substance undergoes chemical changes with change of concentration.

(b) **The Law of Photochemical Equivalence.** Photochemical reactions, like chemical reactions in general, will depend on an "activation" of the molecules taking part. It is believed that the *primary* photochemical reaction consists entirely of activation of this kind, and that *secondary* reactions taking place between activated or partly activated molecules will follow the ordinary kinetic laws.

The activation caused by radiation is governed by the law of photochemical equivalence (Einstein), according to which the radiation energy absorbed by the molecule constitutes a "quantum" of radiation, $h\nu$, where h is Planck's constant [III. (81)], and ν the frequency of the absorbed radiation. The process of activation thus consists in the quantised absorption of energy, given by

$$E = N h \nu \quad (44)$$

for the activation of N molecules.

The activation of a molecule requires a certain minimum amount of energy, depending on the nature of the reaction. This means that a photochemical reaction will be characterised by a certain threshold value for the frequency of the light used, so that light of a lower frequency (*i.e.*, a longer wavelength) will have no effect upon the reaction. The frequency ν is related to the wavelength λ expressed in Ångstrom units by equation X. (54),

$$\lambda\nu = 3.00 \times 10^{18}, \quad (45)$$

so that the activation energy per gram-molecule is given by

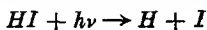
$$E = N_0 h \nu = 3.00 \frac{N_0 h}{\lambda} \cdot 10^{18} \text{ Erg},$$

where $N_0 = 6.06 \times 10^{23}$ and $h = 6.55 \times 10^{-27}$. If E is expressed in calories, this gives for the maximum wavelength which will be effective,

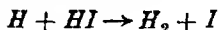
$$\lambda = \frac{2.84}{E} \cdot 10^8. \quad (46)$$

In agreement with this treatment, it is found that light of short wavelength is in general more effective than light of long wavelength. Red light has a wavelength of about 7,000 Å, corresponding to an energy of about 40,000 cal.; thus it is clear that considerable activation energies may occur in photochemical reactions. On the other hand, when once the threshold frequency has been reached (44) shows that N , the number of molecules activated by an amount of radiation E , will decrease with increasing frequency (*i.e.*, decreasing wavelength) of the light absorbed.

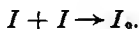
The quantity N in equation (44), which is identical with the number of light quanta absorbed, is obtained by measuring the light absorption E according to the principles described in [a.], ν being known from equation (45) if the wavelength of the light used is known. By comparing N with the number of molecules which have undergone chemical reaction, the "quantum efficiency" of the reaction is obtained. In simple cases these numbers are either identical or bear a simple relation to one another. Thus in the photochemical decomposition of hydrogen iodide with light between about 2,000 and 3,000 Å, it is found that the absorption of one quantum of light leads to the decomposition of two molecules of hydrogen halide. This can be accounted for by assuming that the primary reaction caused by the absorption of the quantum $h\nu$ is



followed by the secondary reactions



and

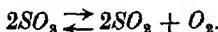


In such cases a direct verification of the law of photochemical equivalence is obtained. In other cases very different behaviour is met with, and the quantum efficiency may be either very much greater or very much smaller than unity. Values greater than unity are usually due to chain reactions, in which the primary activation of one molecule may lead to the reaction of a large number, while the deactivation of primary activated molecules by collision, before reaction has taken place, may reduce the efficiency to an unlimited extent.

The combination of hydrogen and chlorine in the gas phase has already been mentioned as an example of a chain reaction. In this case the effect of the light is to start a reaction chain by photochemical activation, presumably by splitting up the chlorine into two atoms, after which the process goes on as previously described [1. i].

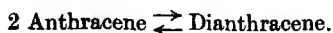
(c) **Photochemical Equilibrium.** If a reversible process is sensitive to light in one or both directions, then under constant conditions of illumination it will be characterised by an "equilibrium state" differing from the equilibrium state in the dark. This "photo-equilibrium" may be regarded as a stationary state, depending on the illumination as well as on the other factors which normally influence an equilibrium. The difference between the photo-equilibrium and the dark equilibrium depends on the independent effect of the light on the two opposed reactions. It thus appears that the similarity between photochemical action and ordinary catalytic effects is only a superficial one.

As an example of a light-sensitive gas reaction of this type we may mention the dissociation of sulphur trioxide,



At 45° in the dark this reaction proceeds immeasurably slowly, the equilibrium corresponding to practically speaking undissociated SO_3 . Under constant illumination by ultra-violet light, a stationary equilibrium is set up which follows the law of mass action, and in which a considerable proportion of the system is present in the form of the dissociation products.

Photochemical equilibrium in solution has been investigated for the equilibrium



On illuminating with ultra-violet light, anthracene is directly converted into dianthracene until a stationary equilibrium state, dependent on the illumination, is reached. In the dark the process is completely reversed, monomeric anthracene being the stable form under these conditions. The photochemical polymerisation obeys Einstein's law of photochemical equivalence.

CHAPTER IX

ELECTROCHEMISTRY

1. ELECTROLYTIC CONDUCTION

(a) **The Conditions for Electrolytic Conduction.** The power of electrical conduction is associated with the motion of electrically charged particles. In "metallic conduction" (which occurs in solid and liquid metals, and in a few metalloids and chemical compounds) it is believed that the conductivity is due to the motion of free electrons. This type of conduction is therefore not accompanied by the transport of matter in the ordinary sense. It is also characterised by the fact that it increases with decreasing temperature, and appears to attain extremely great values in the neighbourhood of absolute zero ("super-conductivity").

Conduction of a different type arises when the moving particles are ions, *i.e.*, atoms or molecules which have acquired electric charges by the gain or loss of electrons. On account of the movement of these ions, this type of conduction is invariably accompanied by the transport of matter, since in an electric field the positive ions will move in the direction of the applied potential, and the negative ions in the reverse direction.

By the action of cathode rays, ultra-violet light and other agencies, gases can be ionised and hence rendered electrically conducting. The conductivity thus produced is however of a transitory nature, since the oppositely charged gas ions will soon discharge one another, thus causing the conductivity to disappear. High temperatures alone are often sufficient to produce ionisation, *e.g.*, incandescent metals may emit electrons [X. 2. c.]. The conductivity present in flames is of a similar nature, being produced by chemical reactions in the flame.

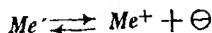
In contrast to these phenomena, *electrolytic conductivity* is a permanent property of the conducting medium. The presence of electrolytic ions thus corresponds to an *equilibrium state* in the medium. Conductivity of this type is usually present only to a small extent in pure substances. It is true that in the case of strong electrolytes the ions necessary for the transport of electricity are often present in the pure substance ; thus, as previously mentioned,

a substance like potassium chloride is completely ionised in the pure state. However, the ionic nature of such substances usually causes them to be solid under ordinary conditions, the ions being fixed in place in an ionic lattice. The conductivity at ordinary temperatures is therefore small, even in the case of strong electrolytes.

On fusion the lattice structure is broken down, and we therefore find high conductivities for fused salts. The same thing takes place when the salt is dissolved in a solvent, and if the dielectric constant of the solvent is high enough to prevent all the ions associating together to form uncharged molecules, then electrolytic conductivity will result.

In other cases, particularly in solutions of acids and bases, the formation of the ions necessary for conduction only takes place by a chemical reaction with the solvent.

(b) **Electrochemical Reactions. Faraday's Law.** The passage of a current through a completely homogeneous conductor will not produce any chemical changes. This will not however apply to the boundary between two different conductors, since here there will not be symmetry with respect to the supply and removal of the moving particles. If one of the conductors is metallic and the other electrolytic, a change will take place at the boundary which may be termed an *electrochemical reaction*. Such a reaction is characterised by the part played by electrons in the electrode process; according to the direction of the current electrons will be either given up to the metal electrode from the electrolytic conductor, or given up to substances present in the electrolytic conductor from the metal electrode. In the simplest case the reaction consists of the solution or deposition of a metal, according to the scheme



Electrochemical reactions are governed by *Faraday's law*, according to which the amount of substance transformed at the electrode, measured in gram-equivalents, is proportional to the quantity of electricity and is independent of the chemical nature of the substance. The amount of electricity required to liberate or dissolve one gram equivalent is

$$F = 96,500 \text{ coulombs,}$$

and is termed a *Faraday*. Expressed in ampere hours, $F = 26.82$. This effect of the current can be used to measure the current or the quantity of electricity, *e.g.*, in the silver voltameter, where it is easily shown that 1 gram of silver corresponds to 894.4 coulombs, the atomic weight of silver being 107.88.

If an aqueous solution is electrolysed, the products at the *negative electrode or cathode* (i.e., the electrode by which the current leaves the apparatus) will be hydrogen or metals, or products produced from these substances by chemical reaction. Similarly, at the *positive electrode or anode* the products will be oxygen, or other electronegative substances, or their reaction products. The components of the electrolyte are the cation, which is discharged at the cathode and has a positive charge, and the anion, which is discharged at the anode and has a negative charge. The amount of charge on one gram equivalent of an ion is F .

The validity of Faraday's law is a direct consequence of the atomic nature of electricity and the resulting view of the ionic charge. As previously stated [III. 1. a.] the elementary quantum of electricity is

$$e = 4.774 \text{ electrostatic units,}$$

or, since one coulomb is equal to 3×10^9 of these units,

$$e = 1.591 \times 10^{-19} \text{ coulombs.}$$

By combining this value with Faraday's constant, we obtain for N_0 , the actual number of molecules in a gram-molecule,

$$N_0 = \frac{F}{e} = 6.06 \times 10^{23}.$$

(c) **Electrolytic Transport.** We will consider a solution containing a completely dissociated salt. When a current passes through this solution the ions will move through the liquid, the cations towards the cathode and the anions towards the anode. In sufficiently dilute solution, the ions move independently of one another, and for a given ion and given conditions the velocity is constant and directly proportional to the applied force, i.e., to the electrical field strength between the electrodes. This direct proportionality is due to the relatively high resistance offered by the liquid to the motion of the ions. The resistance does, however, depend on the nature of the ion, so that the anion and the cation will not in general move with equal velocities when acted on by a given force. The transport of electricity will therefore be shared unequally between the ions. Let u and v be the mobilities of the cation and anion respectively, i.e., their velocities when acted on by unit force. Then for the passage of F units of electricity through a uni-univalent electrolyte the fractions of a gram-ion which pass a given cross-section of the solution will be

$$\frac{u}{u+v} = n \text{ and } \frac{v}{u+v} = 1 - n$$

respectively, and the same ratios will represent the fractions of the total amount of electricity carried by the two ions. n and $1 - n$ are termed the *transport numbers* of the ions (Hittorf), and for dilute solutions are independent of the current carried. The mobilities u and v are also the determining factors in the free diffusion of salts [XI. 2. d.].

In considering further the independent migration of ions we shall take as an example the electrolysis of a solution of silver nitrate between two silver electrodes. We shall imagine the liquid to be divided by a porous partition into the cathode compartment C and the anode compartment A , and consider the passage of F units of electricity from A to C . In A , one equivalent of silver will go into solution in the form of silver ions. $\frac{u}{u+v}$ equivalents of silver will pass from A to C as positive ions, while $\frac{v}{u+v}$ equivalents of nitrate pass from C to A . The number of equivalents of silver added to A will therefore be $1 - \frac{u}{u+v} = \frac{v}{u+v}$: this is equal to the number of equivalents of nitrate ions added, so that the amount of silver nitrate in this compartment will be increased by $\frac{v}{u+v}$ equivalents of salt.

In C , one equivalent of silver ions will be removed, being deposited on the electrode in the form of metallic silver. $\frac{u}{u+v}$ equivalents of silver ions are supplied from A , and $\frac{v}{u+v}$ equivalents of nitrate migrate in the reverse direction. This leads altogether to a decrease of $\frac{v}{u+v}$ equivalents of silver nitrate. The effect of the passage of the current on the composition of the solution is thus to transport $\frac{v}{u+v}$ equivalents of salt from C to A . If ΔC and ΔA represent the increases in the amount of salt in C and A respectively, then

$$\Delta A = -\Delta C = \frac{v}{u+v} = 1 - n.$$

The transport numbers can thus be obtained directly by determining the quantities ΔA or ΔC .

If a solution of hydrogen chloride is electrolysed between platinum electrodes, so that hydrogen is evolved at the cathode and chlorine at the anode, then F units of electricity will liberate one gram equivalent of each substance. At the same time $\frac{u}{u+v} = n$ equivalents of hydrogen will migrate from A to C , and $\frac{v}{u+v} = 1-n$ equivalents of chlorine in the opposite direction. We thus have

$$\Delta A = -Cl - \frac{u}{u+v} H + \frac{v}{u+v} Cl = -\frac{u}{u+v} HCl,$$

$$\Delta C = -H - \frac{v}{u+v} Cl + \frac{u}{u+v} H = -\frac{v}{u+v} HCl,$$

whence

$$\frac{-\Delta A}{-\Delta C} = \frac{u}{v} = \frac{n}{1-n}.$$

The ratio of the transport numbers, or between the velocities of the two ions, is thus equal to the ratio between the amounts of hydrogen chloride lost at the cathode and at the anode, so that n and $1-n$ can also be determined experimentally in this case.

Secondary electrode reactions will not in general prevent the experimental determination of transport numbers, as examination will show. The following table gives the transport number of the cation in 0.1 molar solutions of a few salts at 25°.

TABLE I. *Transport Numbers of Cations*

Salt.	n .
<i>HCl</i>	0.831
<i>LiCl</i>	0.317
<i>NaCl</i>	0.385
<i>KCl</i>	0.490
<i>KBr</i>	0.484
<i>KNO₃</i>	0.513
<i>NH₄Cl</i>	0.490
<i>AgNO₃</i>	0.468

(d) **Specific, Molar and Equivalent Conductivity.** The specific conductivity of a solution is defined as the reciprocal of the resistance

of a column of liquid 1 cm. long and 1 sq. cm. in cross-section. It is denoted by κ . If the solution is placed in a vessel with two fixed electrodes, the resistance is determined by the liquid between the electrodes and will be inversely proportional to the specific conductivity. We thus have

$$\kappa = \frac{k}{r},$$

where r is the measured resistance and k a constant characteristic of the vessel and electrodes. k is known as the "cell constant," and can be measured by using a liquid of known specific conductivity.

The numerical values of the specific conductivity of electrolytes (which will be expressed in reciprocal ohms) are small compared with the corresponding values for metals. Thus for a normal solution of potassium chloride at 18°, $\kappa = 0.0983$, while for copper, $\kappa = 570,000$. This difference is due to the great resistance encountered by ions when passing through a liquid, relative to the resistance to the motion of electrons in the metal.

The *molar conductivity* is defined as the reciprocal of the resistance of that quantity of solution which contains one gram-molecule of solute, when measured with electrodes which are 1 cm. apart. On the basis of this definition it is easy to find the relation between the molar conductivity Λ and the specific conductivity κ , since the conductivities must be in the same ratio as the volumes of liquid between the electrodes. This gives

$$\Lambda = 1000V\kappa = \frac{1000}{m}\kappa \quad (1)$$

where V is the molar volume expressed in litres and in the molar concentration.

The *equivalent conductivity* λ is defined as the conductivity of that quantity of solution which contains one gram-equivalent of solute, measured in the same way as above. It is related to κ by an equation analogous to (1), *i.e.*,

$$\lambda = \frac{1000}{c}\kappa \quad (2)$$

where c is the equivalent concentration. For uni-univalent salts Λ and λ are of course identical.

(e) **Strong Electrolytes.** The following table gives the molar conductivities of a series of salts at 18° at concentrations from 1 to 0.0001 molar.

TABLE II. *Molar Conductivities of Salts*

c	0	0.0001	0.001	0.01	0.1	1
<i>KCl</i>	129.9	129.1	127.3	122.4	112.0	98.3
<i>NaCl</i>	108.9	108.1	106.5	102.0	92.0	74.3
<i>HCl</i>	—	—	377	370	351	301
<i>NaOH</i>	—	—	208	100	183	160
<i>AgNO₃</i>	115.8	115.0	113.2	107.8	94.3	67.8

According to the classical theory of electrolytes [VII. 1.], the considerable increase in λ with increasing dilution is attributed to an increase in the degree of dissociation. According to modern views, however, salts are in general dissociated to such a great extent that this interpretation of the change of conductivity is not permissible. On the other hand, the electrical forces, which have been shown to determine the thermodynamic behaviour of salt solutions, can also be invoked to explain the effect of concentration upon electrolytic conductivity.

It has been previously shown [VII. 2. c.] that the ions in a solution will tend to become surrounded by ions of the opposite sign, so that the change on the "ionic atmosphere" surrounding an ion is opposite in sign to the charge on the ion. When an ion moves under the influence of an electric field, the readjustment of the ionic atmosphere will require a certain time, known as the *relaxation time*: hence the distribution of charge will be symmetrical about the ion (as in the case of a stationary ion), but will be less in front of the moving ion as greater behind it. Since the ion and its atmosphere have opposite charges, this will lead to an electrostatic force, the relaxation force, which will retard the motion of the ion. This force will increase with increasing salt concentration.

In addition to this effect, the motion will be retarded owing to the fact that the ions carry with them to a certain extent the surrounding solvent. On account of the opposite charge of the ionic atmosphere, this means that the ions are moving through a medium which is on the whole moving in the opposite direction. This effect also increases with increasing salt concentration. The effect of concentration on the conductivity is thus explicable in terms of modern views on the dissociation of salts.

Just as in the activity theory, the difference between the conductivities of solutions of different concentrations is found to depend upon the square root of the salt concentrations. If λ is the conductivity at a given concentration c , and λ_∞ the conductivity at infinite dilution, then measurements in dilute solutions can be represented by the equation

$$\lambda = \lambda_\infty - \alpha\sqrt{c}, \quad (3)$$

where α is a constant depending on the salt and on the solvent. Moreover, an equation of this form can be derived from the Debye-Hückel ionic atmosphere theory, using the conceptions outlined above. This theory predicts the value of the constant α in terms of the valencies of the ions, λ_∞ , the temperature, and the dielectric constant and viscosity of the medium. For uni-univalent salts in water at 18°, the theoretical equation is

$$\lambda = \lambda_\infty - (50.5 + 0.225\lambda_\infty)\sqrt{c}. \quad (4)$$

This equation is only valid as a limiting law in the dilute range, but usually permits of extrapolation to obtain λ_∞ , the conductivity at infinite dilution. According to the law of the independent migration of ions, this value must be equal to the sum of the conductivities (or mobilities) of the separate ions at infinite dilution, i.e.,

$$\lambda_\infty = \lambda_{c+} + \lambda_{A-}. \quad (5)$$

Considerable deviations from the limiting law are attributed to incomplete dissociation, or "ion association," as this phenomenon is also termed when dealing with typical salts.

In the last section it was shown that transport experiments could be used to determine the ratio of the mobilities u and v for the two ions constituting a salt. Since the conductivities of ions bearing the same number of charges are in the same ratio as their mobilities, we can write

$$\frac{\lambda_{c+}}{\lambda_{A-}} = \frac{u}{v} = \frac{n}{1-n}. \quad (6)$$

The last two equations show that both the sum and the ratio of λ_{c+} and λ_{A-} are accessible to measurement, so that the conductivities of the individual ions can also be obtained. The following table gives values for some common ions at 18° at high dilutions.

TABLE III. *Ionic Conductivities*

Cations.	Anions.
K^+ 64.0	Cl^- 65.9
NH_4^+ 64.2	Br^- 67.6
Na^+ 43.0	I^- 66.7
Li^+ 33.4	NO_3^- 61.7
Ag^+ 54.1	CH_3COO^- .. 33.7
H_3O^+ 315	OH^- 174

If direct current is used for measuring electrolytic conductivities, then unless special precautions are taken polarisation at the electrodes will interfere with the measurements. For this reason an alternating current is generally used for such measurements, and it is usually found that the frequency of the current is without effect. We should however expect such an effect if the frequency becomes so great that the period of an oscillation is of the same order of magnitude as the relaxation time, since in this event the retarding effect described above will be weakened. In agreement with this prediction, a sensible increase in conductivity has been observed at very high frequencies.

The values of λ_{∞} for different ions in a given medium must depend on the viscous resistance offered to the motion of the ion, and would therefore be expected to decrease as the ionic radius increases. In dealing with this effect it must be remembered that ions are in general *solvated* to some extent, *i.e.*, they are combined with solvent molecules to give more or less permanent complexes. The value of the conductivity is therefore not directly determined by the size of the unsolvated ion: thus in Table III it is not unreasonable that K^+ should have a greater conductivity than Na^+ , since there is independent evidence to show that the sodium ion is more strongly hydrated than the potassium ion.

The very high mobilities exhibited by the H_3O^+ and OH^- ions in aqueous solution may be accounted for by assuming that the observed conductivity is not only caused by the motion of individual ions, but that the mobile protolytic equilibrium between these ions and the H_2O molecules leads to a transfer of the mobile protons from one molecule to another in the direction of the potential gradient.

It is possible to calculate the absolute velocity with which an ion moves in an electric field. The specific conductivity of an ion is equal to the charge carried by the ion in 1 c.c. of solution multiplied by the absolute velocity of the ion, h , in a potential gradient of 1 volt per cm., giving

$$\kappa = \frac{mze}{1000} N_0 h,$$

whence

$$h = \frac{1000\kappa}{mzeN_0} = \frac{\Lambda}{zF}.$$

For the hydrogen ion, $\Lambda = 315$, $z = 1$, $F = 96,500$, giving

$$h = 0.0033 \text{ cm./sec.}$$

while for the potassium ion

$$h = 0.00066 \text{ cm./sec.}$$

These values agree in order of magnitude with the values calculated from Stokes' law [XI. 2. d.].

(f) **Weak Electrolytes.** Solutions of weak acids and bases will conduct electricity to a considerably smaller extent than solutions of strong electrolytes. On dilution their conductivity will change by a much greater proportion than is the case for strong electrolytes, since in addition to the interionic forces there will be a change in the degree of dissociation. The second of these two effects is in general by far the most important. We should not therefore anticipate that equation (3) would be of any use in practice as a limiting law for weak electrolytes.

As mentioned in [VII. 3. a.], the use of the conductivity method for determining the degree of dissociation of weak electrolytes demands a knowledge of the conductivity at infinite dilution. It is not in general possible to obtain this by extrapolation, since a weak electrolyte generally becomes approximately completely dissociated at such a high dilution that measurements can no longer be carried out accurately. For example, for acetic acid we have $K = 1.8 \times 10^{-5}$, and VII. (60) gives $\alpha = 0.99$ at $c = 1.8 \times 10^{-7}$; hence even if VII. (60) could be applied to this very low concentration without taking into account the dissociation of the water itself, there would still remain 1% of undissociated acid.

The value of λ_{∞} is therefore determined by a different method, making use of the law of the independence of ionic mobilities. For acetic acid, EH , we have

$$\lambda_{\infty} = \lambda_{E^{-}} + \lambda_{H_3O^{+}},$$

and the mobilities on the right-hand side can be obtained from data for strong electrolytes, *e.g.*, sodium acetate and hydrogen chloride, using the methods previously described for obtaining the individual ionic mobilities. Or use can be made of the equation $\lambda_{EH} = \lambda_{HCl} + \lambda_{NaE} - \lambda_{NaCl}$. In this connection, however, it should be noted that the anions of weak acids are bases, and their protolysis ("hydrolysis") can have an appreciable effect upon the conductivity, especially at high dilutions.

It is possible to use conductivity measurements for following the course of chemical reactions which are accompanied by a change in the ionic concentration or by the replacement of ions by others of different conductivity. This is the principle on which conductometric titrations are based. On account of the high mobilities of the H_3O^{+} and OH^{-} ions, when a strong acid is added to a solution of

a strong base (or *vice versa*), the point of equivalence is the point at which the conductivity passes through a minimum.

(g) **The Conductivity of Pure Water.** Ordinary distilled water has a specific conductivity of about 10^{-8} . The greater part of this conductivity is due to impurities, chiefly carbon dioxide and to a lesser extent ammonia and substances derived from contact with glass. On very careful purification the conductivity falls to

$$\kappa = 4.0 \times 10^{-8}$$

at 18° C. It increases by about 6% for every degree rise in temperature.

As previously described, the conductivity of pure water is due to its ampholytic nature and the consequent protolytic reaction [VII. (72)]. It is possible to determine the dissociation constant or auto-protolysis constant of water from its conductivity. If we apply equation (1),

$$\Lambda = \kappa \frac{1000}{m},$$

to pure water, then m is the molarity of the dissociated water molecules and Λ the sum of the conductivities of the hydroxonium and hydroxyl ions, *i.e.*,

$$\Lambda_{H_3O^+} + \Lambda_{OH^-} = \kappa \frac{1000}{m}.$$

Inserting the values from Table III, and putting $\kappa = 40 \times 10^{-8}$, we have for the normality of the ions,

$$m = 8.2 \times 10^{-8},$$

and for the dissociation constant of water,

$$K_{(H_2O)} = 6.7 \times 10^{-15},$$

in good agreement with values found by other methods.

In accurate work on the conductivity of aqueous solutions, allowance must be made for the solvent conductivity, which is normally chiefly due to impurities contained in the water. It is however difficult to apply this correction to solutions of acids or bases on account of the protolytic nature of the chief impurities.

(h) **The Effect of the Solvent.** It has already been mentioned that the resistance offered to the motion of a particle increases with the viscosity of the liquid [II. 2. b.]. This factor is of prime importance in considering the dependence of the conductivity upon the medium. For a large group of solvents it is found that the conductivity of a large ion (which is not solvated to a great extent) is at least roughly inversely proportional to the viscosity, so that we can write

$$\Lambda \eta \sim \text{constant}$$

where the value of the constant varies between 0.5 and 1 when the viscosity is expressed in absolute units.

The conductivity of a salt also depends upon the "dissociating power" of the solvent. As previously stated [VII. 2. a.], this property is related to the dielectric constant, since according to Coulomb's law the forces between charged particles are inversely proportional to this constant.

TABLE IV. *Dielectric Constants of Liquids*

Hydrogen cyanide	96
Formamide	> 84
Water	81
Formic acid	58
Nitromethane	39
Acetonitrile	36
Methyl alcohol	33
Ethyl alcohol	26
Acetone	22
Sulphur dioxide	16
Pyridine	12
Phenol	10
Ether	4.8
Chloroform	4.6
Benzene	2.3
Hexano	1.9

If we consider the conductivity of a salt in the above series of solvents, values are found which on the whole decrease parallel with the dielectric constants. There are, however, numerous exceptions, *e.g.*, hydrogen cyanide is not nearly such a strongly dissociating medium as would be expected from its position in the series.

The effect of dilution upon the conductivity is usually much less simple and predictable in solvents of low dielectric constant than for aqueous solutions.

2. ELECTRICAL POTENTIAL AND ELECTROMOTIVE FORCE

(a) **Potential Differences between Two Phases.** The **Electrochemical Potential**. The difference in electrical potential between two points is defined in electrostatic theory as the work done by electrical forces in transferring unit quantity of electricity from one point to the other. A corresponding definition can also be applied to chemical systems, provided that the two points considered are in the same medium. In particular, if there is electrical equili-

brum in a homogeneous phase which is a conductor of electricity, the electrical potential may be said to be the same throughout.

If on the other hand the chemical system being considered is not homogeneous, but consists for example of two different phases, then it is easily seen that the above definition can no longer be used directly. According to the definition, the potential difference is measured by the amount of work done in transferring unit quantity of electricity from one phase to another. An abstract "quantity of electricity," however, exists only in electrostatic theory, and not in actual systems, where every electric charge is associated with a *particle* (ion or electron), the chemical nature of which will play a part in determining the work of transference. In other words, for a given charged particle this work will depend not only on electric forces, but also on chemical ones, *i.e.*, forces which are also effective in the transfer of uncharged particles.

The definition of potential used in electrostatics is therefore not immediately applicable to chemical systems on account of the effect of the change of medium. Conversely, in considering the chemical potential of charged particles a term must be added to allow for the electrical effects, so that the thermodynamic equations must contain an extra variable of electrical origin. Purely thermodynamical methods do not make it possible to distinguish between the effects due to electrical and chemical factors, but we may represent the thermodynamic behaviour of ions by replacing the chemical potential μ by the *electrochemical potential* in the thermodynamic equations. The significance of $\bar{\mu}$ is given by the equations thus obtained, and its value will depend upon state of electric charge, as well as on the surrounding medium and the other variables representing the state of the system. It appears reasonable to attempt to split up the electrochemical potential thus defined into a chemical and an electrical component by means of the expression

$$\bar{\mu} = \mu + \psi z F,$$

where z is the valency of the ion and F represents a Faraday, but it is clear that this formulation is meaningless as long as ψ is only defined by the above electrostatic definition of electrical potential.

The determination of ψ , which we shall term the *electrical potential* of the phase, becomes possible if we arbitrarily define ψ as zero in a phase which is unchanged, *i.e.*, electrically neutral. If we consider an isolated phase satisfying this definition, then the chemical potential μ_1 of any ion in this phase will be given by

$$\mu_1 = \bar{\mu}_1,$$

and since $\bar{\mu}_1$ is thermodynamically defined, this constitutes a definition of μ_1 . Further, since the charge on a phase has practically no effect on its chemical composition, we can define μ_1 as being independent of the state of charge: under these conditions the above equation for $\bar{\mu}$ constitutes a general definition of ψ .

The division of the electrochemical potential into a chemical and an electrical component is thus permissible, and leads in principle to satisfactory definitions of the single components. Special difficulties are however encountered in the experimental determination of these electrical potentials, since electrical potentials of the order of magnitude occurring in chemical systems are produced by immeasurably small quantities of ions, thus preventing the use of ordinary thermodynamic methods for the determination. We shall not however consider here the possibility of actual determinations of electrical potentials and chemical potentials of ions in systems of this kind, since the only point which is essential for our present considerations is the proof that these quantities can be defined and used in principle.

This proof of course also applies to other functions such as the activity and activity coefficient of electrically charged particles. For example, if an ion is in distribution equilibrium between two isolated phases, then if the ionic concentrations are so small that the phases can be regarded as uncharged, these concentrations will be inversely proportional to the absolute activity coefficients of the ions [IV. 3. e.].

In agreement with this, the difference in electrical potential between two phases in which ions are distributed can be attributed to the difference in the absolute activity coefficients of one or more ions in the two phases. Such a difference will be present in general for all ions, and there will therefore be a tendency for non-uniform distribution of ions and hence the production of electric charges on the phases. For example, if metallic zinc is brought into contact with a solution of a zinc salt, it will be possible for zinc ions to pass either from the metal into the solution, or in the reverse direction. In the first case the metal will be charged positively relative to the solution, while this will be reversed in the second case. As a second example we may consider an aqueous solution of NaCl shaken up with ether until equilibrium is reached: there will then be an unequal distribution of the two sorts of ions between the two liquid phases, and hence a displacement of charge. However, an account of the large value of the ionic charges, the deviation from equivalence between positive and negative particles in a given phase will always be too small to be detected analytically.

In the above examples we may regard the potential produced as

due to a competition between chemical and electrical forces. The former are individual in character, and attempt to set up an uneven distribution, while the latter attempt to restore electrical equivalence, in which they are supported by the thermal motion of the molecules. As a result of this there will be an electrical double layer at the interface between the two phases, since the ions which are in excess in the two phases will tend to approach one another as closely as possible, but are held back in their own phase by the chemical forces.

It is possible in principle to calculate by ordinary electrostatic principles the difference in potential due to this unequal distribution, since it is equal to the work needed to take a unit electric charge through the double layer. It is however necessary to know both the electrical density and the thickness of the layer in order to carry out this calculation. It should also be remembered that each half of the layer cannot be regarded merely as a plane, since on account of thermal agitation they will be diffuse in character.

The mechanism described here for the setting up of a potential difference between two phases can also be applied to a contact between two metals. As previously mentioned, the electrons constitute the mobile components of metals, and we may assume that they have a tendency to become unequally distributed between the two metals, thus setting up an electrical double layer and a difference of potential. This is termed the *contact potential* $\psi_{1,2}$, and is equal to the difference between ψ_1 and ψ_2 , the potentials of the two metals. Systems composed of metals do not therefore differ from other systems in this respect. The interface between two metals does however exhibit very much simpler behaviour than the interfaces between other phases. This is because in the case of metals the transport of electricity when a current is passed is due entirely to one kind of particle, the electron: since electrons cannot accumulate, the current therefore brings about no changes whatever at the point of contact. This simplification is entirely absent at interfaces between metals and electrolytic conductors, since in this case the conducting particles are different on the two sides of the plane of contact, so that the passage of a current must necessarily bring about chemical changes.

(b) **The Electromotive Potential of Metals.** Let us consider two blocks of metal in direct contact. In agreement with what has been said above, there will be an electrical potential difference between the metals: however, this will not be apparent when the two metals are connected to a voltmeter, since no matter what the nature of the metallic connection, the instrument will give no deflection provided the whole system is at a constant temperature. This is

an experimental result, and also follows from the laws of thermodynamics, since the system admits of no possible compensation for the electrical work which would be associated with the passage of current through the system.

A closed circuit composed of metals will thus exhibit no current if kept at constant temperature. If however there are differences in temperature between the different points of contact, this will set up a "thermo-electric force" which can bring about the transformation of thermal potential energy into electrical energy (or *vice versa*) when a current passes through the system.

The absence of current in the isothermal circuit is thus a consequence of energy relations, but may also be interpreted in terms of the state of electrical equilibrium between two metals in direct contact. As mentioned above, this equilibrium depends on the identity of the electrochemical potentials of any mobile particle in the two phases. Since the atoms and ions in the metal are completely fixed in the crystal lattice (and those in a liquid metal may be regarded as motionless compared with the electrons), the electron is the only mobile particle, and the condition for electrical equilibrium between the two phases may be expressed by

$$\bar{\mu}_{\Theta(1)} = \bar{\mu}_{\Theta(2)}. \quad (7)$$

Since in the equilibrium state the system of metals has no effect upon a voltmeter, equation (7) has an important practical bearing on electrical measurements. This may be expressed by introducing a quantity termed the *electromotive potential* π , defined by

$$\pi F = -\bar{\mu}_{\Theta}. \quad (8)$$

The electromotive potential is thus a quantity which is proportional to the electrochemical potential of the electron, but has the opposite sign. It is of great importance in determining the behaviour of the metal towards a measuring instrument. In a system of metals in contact with one another each metal will have its individual electrical potential, but the electromotive potential will be the same throughout the whole system of metals. If one such system is placed in contact with each of the two terminals of an instrument for measuring potential, the deflection will give the difference between the electromotive potentials of the two systems.

The relation between the electrical potential and the electromotive potential of metals is obtained in the following way. As described above, the electrochemical potential of an ion of species 1 can in general be split up into two terms according to the equation

$$\bar{\mu}_1 = \mu_1 + z_1 F \psi, \quad (9)$$

and hence for an electron in a given metal (1),

$$\bar{\mu}_{\Theta(1)} = \mu_{\Theta(1)} - F \psi_1 \quad (10)$$

By introducing the electromotive potential, as defined above, this equation becomes

$$F \pi_1 = -\mu_{\Theta(1)} + F \psi_1. \quad (11)$$

The electron concentration in a given metal is practically a constant quantity even for high charges, since even very small deviations from equivalence correspond to very great charges. Hence μ_{Θ} , the chemical potential of the electron, must be a characteristic constant for a given pure metal at a given temperature. The difference between the electromotive potentials of two metals is therefore expressed by

$$F(\pi_1 - \pi_2) = -(\mu_{\Theta(1)} - \mu_{\Theta(2)}) + F(\psi_1 - \psi_2), \quad (12)$$

where the first term on the right-hand side depends on the nature of the metals. If this equation is applied to two metals in contact, we have $\psi_1 - \psi_2 = \psi_{12}$ = the contact potential, while $\pi_1 = \pi_2$. Hence in general,

$$\mu_{\Theta(1)} - \mu_{\Theta(2)} = F \psi_{12}, \quad (13)$$

and

$$\pi_1 - \pi_2 = (\psi_1 - \psi_2) - \psi_{12}. \quad (14)$$

The difference of electromotive potential between two metals is thus equal to their difference in electrical potential minus the contact potential.

If the two metals are the same, $\psi_{12} = 0$, in which case the difference in electromotive potential is equal to the difference in their electrical potentials, *i.e.*,

$$\pi_1 - \pi_2 = \psi_1 - \psi_2. \quad (15)$$

Equation (13) shows that the contact potential is given by the difference of two quantities $\mu_{\Theta 1}$ and $\mu_{\Theta 2}$, which are characteristic of the single metals and not of the combination. This is in agreement with the law of additivity on which depends the absence of current in a circuit composed of metals.

We shall denote the potential corresponding to μ_{Θ} by ψ^* , defined by

$$\psi^* F = -\mu_{\Theta}. \quad (16)$$

Equation (13) then becomes

$$\psi_1^* - \psi_2^* = -\psi_{1,2} \quad (17)$$

and (11) becomes

$$\pi = \psi + \psi^*. \quad (18)$$

The quantity $-\psi^*$ thus introduced is a constant characteristic of a given metal and is a measure of the tendency of an electron to leave the metal when the latter is electrically neutral.

(c) **Galvanic Cells.** When a number of conductors are connected in series so that the passage of a current causes changes to take place at the boundaries between the conductors, the system is generally termed a *galvanic cell* or *element*. The simplest form of such a cell consists of two metallic electrodes separated by electrolytic conductors. If the system as a whole is not in chemical equilibrium, the two metallic electrodes will possess different electromotive potentials. The difference between the electromotive potentials of the electrodes is termed the *electromotive force* of the cell, and can be read off directly on a suitable instrument for measuring potentials. If the electrodes are connected by a metallic conductor, a current will pass through the cell, and electrical energy will be produced at the expense of other forms of energy.

When a current is being produced, the difference between the electromotive potentials of the electrodes (*i.e.*, the voltage of the cell) will fall below the electromotive force, since we have

$$\pi = (r_i + r_v)i,$$

$$\pi' = r_v i,$$

where π is the electromotive force, π' the voltage of the cell, r_i the internal resistance of the cell, and r_v the external resistance of the circuit carrying the current. If the external resistance is very great, this difference disappears. Further, the voltage may fall on account of chemical changes produced in the cell by the passage of the current.

A galvanic cell may be treated from two essentially different points of view. On the one hand, we may consider the dependence of the single boundary potentials on the nature of the phases in contact. On the other hand, we may regard the whole cell as a chemical system capable of producing electrical energy, which can therefore be treated by ordinary thermodynamic methods. In both cases the object of the treatment is to calculate the electromotive force of the cell, or its dependence upon the variables involved. It is only possible to obtain quantitative results of this kind when the cell functions in a completely reversible manner.

(d) **Electromotive Reversibility.** To explain this concept we will

consider a galvanic cell E connected to another source of current E_0 , the positive and negative poles of E being connected respectively to the positive and negative poles of E_0 . If it is possible to vary the voltage of E_0 continuously, it will be possible to find a particular voltage π for which no current passes through the combination of E and E_0 : this may be detected, for example, by means of a galvanometer inserted in the circuit. If now an increase or decrease of the voltage by a very small amount $\Delta\pi$ causes a current to flow in opposite directions, the first condition for the reversibility of the cell E will be satisfied. In a *reversible galvanic cell* the electromotive force is thus independent of the direction of the current.

If this condition is satisfied, then the other condition for reversibility will also be practically always satisfied, *i.e.*, that the passage of current in two opposite directions through the cell brings about chemical changes differing only in their direction. We can therefore define a reversible galvanic cell as an electromotive combination in which the direction of the current has no effect on the electromotive force, and only affects the direction of the chemical process responsible for the electromotive force. These criteria can only be applied when the current is infinitesimal, since with finite currents the cell is not completely reversible, and a part of the electrical energy which would otherwise be available is lost as heat inside the cell. This "ohmic" heat must not be confused with reversible heat change in the cell, which may be either positive or negative, and which (for a given cell) depends only on the quantity of electricity, and not on the strength of the current.

All transformations in the cell must of course follow Faraday's law, *i.e.*, one equivalent corresponds to a quantity of electricity $F = 96,500$ coulombs. Since electrical energy is expressed as the product of potential difference and quantity of electricity, each equivalent transformed in the cell will produce πF units of electrical energy, where π is the electromotive force of the reversible cell. This quantity must also represent the maximum electrical work which the chemical process in question can produce, since π represents the limiting value of the potential difference which the cell can overcome. If the transformation taking place in the cell is accompanied by a volume change, a certain amount of mechanical work will also be done; this is, however, generally very small compared with the electrical work.

There is an obvious analogy between electromotive reversibility and the mechanical reversibility of other processes, *e.g.*, when a liquid vaporises under the pressure of its saturated vapour, and produces a certain volume increase against an equal opposing pressure. In both cases the maximum work is obtained when the

intensity factor which determines the direction of the process has its maximum value. This agrees with the previous statement that reversible processes take place through an equilibrium state.

(e) **The Thermodynamic Treatment of Galvanic Cells.** The maximum electrical work for a reversible galvanic cell is measured by its electromotive force, and the heat change when the cell functions reversibly can also be determined experimentally. It is therefore easy in principle to test the second law of thermodynamics, *e.g.*, equation I. (11),

$$c = T \left(\frac{dA}{dT} \right)_v,$$

for a cell working at constant volume. Thus if the cell is placed in one calorimeter (I.) and the circuit completed through a resistance placed in a second calorimeter (II.), then if the resistance is sufficiently high, the quantity of heat developed in (II.) is equal to A , and the quantity of heat developed in (I.) is equal to $-q$. If the experiment is carried out at different temperatures, dA/dT can also be measured and the above equation tested. It will be seen that in the above experiment it is not necessary to know the nature of the process taking place in the cell and leading to the production of a current. In practice, however, it would often be difficult to make the cell work completely reversibly and at the same time produce a sufficient current for measuring the heat change.

This difficulty is avoided if the process in the cell is known, so that ΔE for the transformation can be measured by more direct methods. In this case we use equation I. (12),

$$A + \Delta E = T \left(\frac{dA}{dT} \right)_v,$$

or for a transformation involving z equivalents,

$$\Delta E + zF\pi = TzF \left(\frac{\partial \pi}{\partial T} \right)_v. \quad (19)$$

This is Helmholtz's equation for the electromotive force.

The above treatment is only exact when the volume is constant. A result of more general application is obtained from the thermodynamic equations I. (25), where it is easily seen that the electrical work produced by a reversible cell can be identified with the quantity Δ_0 .

If the cell works at constant temperature and pressure, equation I. (31)

$$dG = -SdT + vdp - \Delta_0,$$

gives for the electrical work,

$$- \Delta G = F\pi z. \quad (20)$$

On introducing I. (34) and I. (37), this gives

$$- \Delta V = zF \left(\frac{\partial \pi}{\partial p} \right)_T, \quad (21)$$

and

$$\Delta H + zF\pi = TzF \left(\frac{\partial \pi}{\partial T} \right)_p. \quad (22)$$

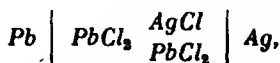
The same result is obtained by writing $\Delta_0 = \pi F dz$ in I. (31) and differentiating the equation thus obtained with respect to T and p , using the usual rules for partial differentiation.

Equation (21) shows that the effect of pressure on the electromotive force is determined by the volume change which accompanies the process producing the current, while (22) relates ΔH , the heat absorption as measured calorimetrically, to the electromotive force and the temperature coefficient of the electromotive force at constant pressure.

In using (22) for numerical calculations, if ΔH is expressed in calories and π in volts, we must put $F = 96,500$ coulombs, and 1 volt-coulomb = 1 joule = 0.239 cal., giving

$$\Delta H + 23060 z\pi = 23060 zT \left(\frac{\partial \pi}{\partial T} \right)_p.$$

We will apply this equation to an example. Consider the cell represented by



i.e., a lead electrode and a silver electrode immersed in a solution which is saturated throughout with lead chloride, and saturated with silver chloride round the silver electrode. At atmospheric pressure the electromotive force is found to be given by the empirical equation

$$\pi = 0.4917 - 0.000165t$$

where t is the centigrade temperature. For $t = 20^\circ$ we thus have

$$\pi = 0.4884,$$

and

$$\left(\frac{\partial \pi}{\partial T} \right)_p = -0.000165.$$

Putting these values in equation (22) and also $z = 2$, we find

$$\Delta H = -24,750 \text{ cal.}$$

The current producing process in this cell is

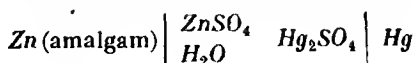


and ΔH , the increase in the heat content, is thus equal to the difference between the heats of formation of the two chlorides. The heat of formation of $PbCl_2$ is 82,770 cal., while for $2AgCl$ it is 58,760 cal. The difference,

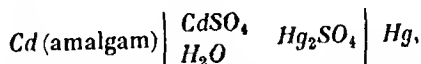
$$\Delta H = -24,010 \text{ cal.}$$

agrees well with the value from the electromotive force.

Cells of the same type as the above are the Clark cell,



and the Weston cell,



both of which are extremely constant and reproducible, and are therefore used as *standard cells* in electrometric measurements. When the solutions are saturated with the soluble hydrated sulphates the electromotive force in volts is given by,

$$\text{Clark cell} - \pi = 1.4328 - 0.0012(t - 15^\circ)$$

$$\text{Weston cell} - \pi = 1.0181 + 0.00004(t - 15^\circ).$$

These equations may be used to calculate the thermal data, as in the above case.

We can conclude in general from equation (22) that the electromotive force will increase with increasing temperature if the electrical energy produced by the cell, $23060z\pi$, is greater than the heat evolved by the process, while it will decrease with increasing temperature if the reverse is the case. If the cell functions reversibly, which can be realised in practice by completing the circuit with a very high external resistance, then the heat produced in this resistance will correspond to the maximum electrical work of the process. If this amount of work is greater than the total energy change, then when the cell is functioning isothermally and reversibly it will absorb heat from the surroundings, analogous to the heat absorbed in the reversible vaporisation of a liquid. The question of how far this heat can be said to be converted into electrical energy has been discussed in detail in the thermodynamic introduction [I. 2. a.].

It was at one time assumed that the electromotive force was directly determined by the heat evolved in the process (William Thomson, Lord Kelvin), in agreement with Berthelot's views on affinity, according to which the heat of reaction was supposed to be a direct measure of the affinity. This principle is approximately fulfilled in many cases, and equation (22) shows that it will be exactly valid if the temperature coefficient of the electromotive force is zero. In other cases, however, it fails completely, *e.g.*, for the concentration cells described in the next paragraph.

(f) **Concentration Cells.** In certain galvanic combinations the process responsible for the electromotive force consists only of a *concentration change*. ΔG for processes of this kind can usually be determined by some other method, and we can then use (20) to predict the electromotive force of this type of cell. Usually we shall be dealing with a differential concentration change, *i.e.*, the transfer of a substance from one definite concentration to another. For this type of process I. (47) gives

$$\Delta G = \mu_{(2)} - \mu_{(1)},$$

there $\mu_{(1)}$ and $\mu_{(2)}$ are the chemical potentials of the substance transferred in the initial and final states respectively. Introducing (20), we have

$$F\pi z = \mu_{(1)} - \mu_{(2)}, \quad (23)$$

whence by using IV. (59),

$$F\pi z = RT \ln \frac{a_{(1)}}{a_{(2)}}, \quad (24)$$

or, if the gas laws hold,

$$F\pi z = RT \ln \frac{x_1}{x_2} = RT \ln \frac{c_1}{c_2}. \quad (25)$$

Since (24) and (25) refer to one gram-molecule, z is the number of electrical equivalents contained in one gram-molecule.

Putting $R = 1.98$, converting to decadic logarithms and expressing the energy in calories [VI. 4. b.], this becomes

$$4.57 T \log \frac{c_1}{c_2} = 23060 z\pi$$

$$\pi = 0.0001982 \frac{1}{z} T \log \frac{c_1}{c_2}, \quad (26)$$

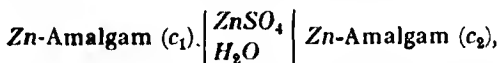
giving for $T = 20^\circ$,

$$\pi = 0.0581 \frac{1}{z} \log \frac{c_1}{c_2} \quad (27)$$

These equations can be used to calculate π , z or $\frac{c_1}{c_2}$ when the other two quantities are known. The agreement between the observed and calculated values then constitutes a test of the validity of the gas laws for the system considered.

For concentration cells obeying (25), ΔG and the electrical work are directly proportional to the temperature, and hence from I. (37) $\Delta H = 0$. When the cell is working reversibly at constant volume it will therefore absorb an amount of heat equivalent to the work done, just as in the expansion of an ideal gas.

A simple example of this type of cell is provided by *amalgam cells*, in which the electromotive process is the transfer of a metal dissolved in mercury from one concentration to another. For example, if a cell is constructed with electrodes consisting of two zinc amalgams of concentrations c_1 and c_2 and zinc sulphate solution as electrolyte, *i.e.*,



the electromotive process will consist in solution of zinc at the negative pole and deposition of zinc at the positive pole, accompanied by displacement of zinc sulphate from the liquid round the positive pole to the liquid round the negative pole. Since the concentration of the zinc sulphate is the same throughout the solution, this displacement will have no effect on the electromotive force, so that in calculating the latter we can assume that the transfer of metallic zinc from one amalgam to the other is the electromotive process. If we introduce the values of c_1 and c_2 in equation (26), then the values of π for dilute amalgams agree with experiment if the value 2 is used for z . This shows on the one hand that the gas laws are valid for dilute zinc amalgams, and on the other hand that the molecules formed when zinc is dissolved in mercury contain one zinc atom. The last conclusion is generally valid for amalgams (and also for vapours of metals) so that equation (26) can often be applied to amalgam concentration cells when z is put equal to the valency of the metal.

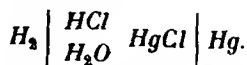
A second type of concentration cell can be set up with hydrogen electrodes. If hydrogen is passed over a platinum plate electrolytically platinised, an electrode is obtained which behaves essentially as a *hydrogen electrode*. If two such electrodes in contact

with hydrogen at different pressures p_1 and p_2 are immersed in the same solution (e.g., hydrochloric acid) a cell is obtained in which the electromotive process consists entirely in the transference of hydrogen from one pressure to another. Since $z = 2$, the electromotive force is given by

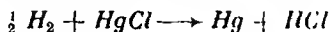
$$\pi = 0.0000991 T \log \frac{p_1}{p_2}$$

by analogy with the above.

Particular importance attaches to concentration cells in which the electromotive process is the transfer of a solute from one aqueous solution to another. In the cell



where one electrode is a hydrogen electrode and the other a "calomel electrode" [i. o.], the electromotive process is



If two such cells with different HCl concentrations are placed in opposition to one another, then if the hydrogen pressure is the same in the two cells, the electromotive process will consist in the transfer of HCl from one concentration to another, since the simultaneous transfer of H_2 , $HgCl$ and Hg can have no effect upon the electromotive force.

According to equation (24) the electromotive force of the combination is given by

$$F\pi z = RT \ln \frac{a_{HCl(1)}}{a_{HCl(2)}}$$

or, using VII. (5), VII. (8) and VII. (9),

$$F\pi z = RT \ln \frac{(a_{H^+} + a_{Cl^-})_{(1)}}{(a_{H^+} + a_{Cl^-})_{(2)}}$$

or,

$$F\pi z = 2 RT \ln \frac{c_1}{c_2} + 2 RT \ln \frac{f_{\pm(1)}}{f_{\pm(2)}}. \quad (28)$$

In extremely dilute solutions we can put $f_{\pm} = 1$, and the electromotive force of the cell can be calculated from the gas laws, the last term in (28) disappearing. Further, as already mentioned [VII. 2. e.], measurements of this kind offer a possible method for determining the activity coefficients of ions, since equation (28) shows that f_{\pm} is directly related to π . Electrometric measurements

can in many cases be carried out with high accuracy, and the method has therefore been of considerable importance in the development of the theory of strong electrolytes.

In the case of concentration cells involving concentrated solutions of the salt transferred, the electromotive force can always be calculated if the activity coefficients are known. The following method of calculation (due to Helmholtz) depends on this principle, but is formally somewhat different.

The chemical potentials μ_1 and μ_2 of the salt and the solvent in a salt solution are related by the Gibbs-Duhem equation IV. (48),

$$x d\mu_1 + (1-x) d\mu_2 = 0.$$

By integrating this equation between the concentration limits x_1 and x_2 we obtain

$$\mu_{1(x_2)} - \mu_{1(x_1)} = \int_{x_1}^{x_2} \frac{1-x}{x} d\mu_2.$$

Introducing IV. (60),

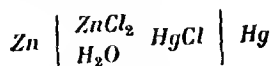
$$d\mu_2 = RT d \ln a_2 = RT d \ln p_2,$$

we see that the integration can be carried out if p_2 , the vapour pressure of the solvent, is known as a function of x in the concentration interval in question. Introduction of (23) gives

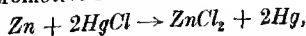
$$Ez = RT \int_{x_1}^{x_2} \frac{1-x}{x} d \ln p_2, \quad (29)$$

from which π can be calculated.

As an example of this calculation we may consider combinations of the cell

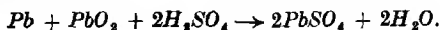


in which the electromotive reaction is



HgCl being present as a solid and ZnCl_2 in solution. If two such cells are opposed to one another, then apart from the transfer of pure solids or liquids from one cell to another (which cannot affect the electromotive force), the electromotive process consists in the transfer of ZnCl_2 from one solution to another. If we introduce the concentrations x_1 and x_2 and express p_2 as a function of x in such a form that the equation can be integrated, we obtain an expression for π which can be directly compared with experiment.

Similar principles can be used for treating more complicated cells, e.g., the lead accumulator. When a lead accumulator is discharging, the process is



When two such cells are opposed to one another, then (by analogy with the above treatment) the net electromotive process will be the transfer of H_2SO_4 from the cell with acid concentration c_1 to the cell with acid concentration c_2 , and the transfer of an equivalent amount of water in the reverse direction. By analogy with (23), we have for this process,

$$F\pi z = \mu_1(x_1) - \mu_1(x_2) - (\mu_2(x_1) - \mu_2(x_2)),$$

where μ_1 and μ_2 are the chemical potentials of sulphuric acid and water respectively, while x_1 refers to the higher and x_2 to the lower acid concentration. For water the difference in chemical potential is given by IV. (74) as

$$\mu_2(x_1) - \mu_2(x_2) = RT \ln \frac{p_2(x_1)}{p_2(x_2)},$$

and a similar expression will hold for the sulphuric acid. The vapour pressure of the latter substance is however so small as to be immeasurable, and we therefore use equation (29) to relate the chemical potential of the sulphuric acid to that of the water. We thus obtain

$$F\pi z = RT \ln \frac{p_2(x_2)}{p_2(x_1)} + RT \int_{x_1}^{x_2} \frac{1-x}{x} d \ln p_2,$$

where it is easily seen that both terms are positive. Further, $z = 1$, since 1 mole of acid and of water is transported per electrochemical equivalent.

The following table gives the values directly measured and those calculated from the above expression.

TABLE V. *E.M.F. of the Lead Accumulator*

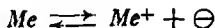
% H_2SO_4	π (exp.)	π (calc.)
64.5	2.36	2.38
52.2	2.25	2.26
35.3	2.10	2.10
21.4	2.01	2.00
5.16	1.89	(1.89)

The value 1.89 for 5.16% H_2SO_4 was used as a basis for the calculated figures in the last column.

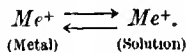
(g) **Electrode Potentials.** At a high temperature in a vacuum, metals will emit electrons and hence become *positively* charged, a

thermal electrostatic equilibrium being set up. On the other hand, in contact with a solvent they will exhibit a tendency (depending on the nature of the metal and of the solvent) to send metal ions into the solution, which will charge the metal *negatively* relative to the solution. As previously described, an electrical double layer and a difference of electrical potential will be set up. This will oppose the tendency of the ions to dissolve, and if no other chemical changes take place the process will come to a standstill before detectable quantities of metal have gone into solution. We then have a state of equilibrium in which the tendency for the ions to dissolve is balanced by the electrical potential produced: this equilibrium is characterised by the identity of the electrochemical potentials of the ions in the two phases.

If the solution only contains the few metal ions produced by the above solution process, the potential will in practice be indeterminate. If on the other hand the solution already contains a given concentration of a salt of the metal in question, the potential will be in principle fixed. For example, if a rod of metallic silver is in contact with a 0.1 molar silver nitrate solution, a well-defined potential will be set up between the metal and the solution. The equilibrium can be expressed by



or, since the solid metal is made up of metal ions and electrons,



The equilibrium between the electrode and the solution is determined by the identity of the electrochemical potentials. If we denote the metallic phase by (1) and the electrolytic phase by (I), application of the general equation (9) to the present case gives

$$\mu_{(1)} + zF\psi_1 = \mu_{(I)} + zF\psi_I \quad (30)$$

or,

$$\mu_{(I)} - \mu_{(1)} = zF(\psi_I - \psi_1) = zF\psi_{1,I} \quad (31)$$

where $\mu_{(I)} - \mu_{(1)}$ is the difference between the chemical potentials of the metal ion in the metal phase and in solution, and $\psi_{1,I}$ is the electrical potential difference between the solution and the metal. It is not possible to calculate $\psi_{1,I}$, since we have already seen that the ionic potentials $\mu_{(I)}$ and $\mu_{(1)}$ cannot be determined by thermodynamic methods. However, as long as we consider dilute solutions in a constant medium it is possible to use equation (31) for calculating the *effect of concentration* on the potential $\psi_{1,I}$ at constant temperature and pressure, since under these conditions the chemical

potentials (as well as the activities and activity coefficients) of the dissolved ions obey the laws described in [VII. 2.].

Of the two chemical potentials $\mu_{(1)}$ is a constant as long as the metal is present in an unchanged condition. $\mu_{(1)}$, on the other hand, depends on the concentration of ions in the solution, being given by IV. (59) as

$$\mu_{(1)} = RT \ln a + k.$$

The effect of concentration upon the potential of the electrode is readily obtained by inserting this equation in (31). This gives

$$\mu_{(1)} - RT \ln a - k = \psi_{1,1} z F,$$

or,

$$\psi_{1,1} = - \frac{RT}{zF} \ln a + \psi_{1,1(0)}, \quad (32)$$

where a is the ionic activity, z the charge on the ion (with the correct sign), and $\psi_{1,1(0)}$ a constant which is equal to the value of the potential at unit ionic activity.

Equation (32) expresses the potential difference $\psi_{1,1} - \psi_{1,1(0)}$ between two electrodes of the same metal immersed in solutions of a salt of this metal with ionic activities a and 1 respectively. According to (15), the difference

$$\psi_{1,1} - \psi_{1,1(0)} = \psi_1 - \psi_{1(0)} - (\psi_1 - \psi_{1(0)})$$

can be written

$$\psi_1 - \psi_{1(0)} - (\pi - \pi_0).$$

We thus see that if the two solutions surrounding the electrodes are brought into contact with one another in such a way that they assume the same electrical potential (which can be realised approximately—*cf.* [n., o.]), then the combination thus formed will have an electromotive force given by

$$\pi_1 - \pi_0 = \frac{RT}{zF} \ln a_1, \quad (33)$$

or, if the solutions constituting the combination have arbitrary ionic activities a_1 and a_2 ,

$$\pi_1 - \pi_2 = \frac{RT}{zF} \ln \frac{a_1}{a_2}. \quad (34)$$

The electromotive potentials occurring in these equations are often termed "electrode potentials," the potential of the solution being taken as zero. The equations express the effect of the ionic activity upon these potentials, and can therefore be used to determine ionic activities from electrometric measurements. In this way it is possible to investigate the solubilities of sparingly soluble salts

and equilibria involving complex ions containing the electrode metal.

If the solution is very dilute, a can be replaced by the ionic molarity m , and we obtain the approximate equation

$$\pi - \pi_0 = \frac{RT}{zF} \ln m. \quad (35)$$

If π is measured in volts and decadic logarithms are used, the factor $\frac{R}{F}$ can be replaced by 0.0001982. At 70° the equation becomes

$$\pi - \pi_0 = \frac{0.0581}{z} \log m. \quad (36)$$

(h) **Standard Electrode Potentials. The Electrochemical Series.** The above considerations deal with the effect of the ionic activity, and will of course give no information about the effect of the *nature* of the metal or the value of the electromotive or electrical potential. This question can only be decided by experiments in which electrodes of different metals are combined to give a galvanic cell. We shall now see what information can be derived from the electromotive force of such combinations.

We will consider solutions of two salts in the same solvent, each

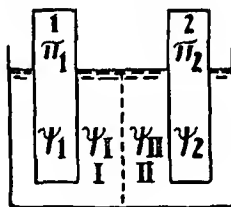


Fig. 1.

containing an electrode of the appropriate metal, the solutions being in contact as shown in Fig. 1. We will denote the metals by 1 and 2, and the corresponding solutions by I and II. The difference between the electrical potentials of the metal electrodes, $\psi_1 - \psi_2$, will obviously be given by the sum of the three contact potentials occurring in the cell,

$$\psi_1 - \psi_2 = \psi_{1,I} + \psi_{I,II} + \psi_{II,2}, \quad (37)$$

since

$$\psi_{1,I} = \psi_1 - \psi_I$$

$$\psi_{I,II} = \psi_I - \psi_{II}$$

$$\psi_{II,2} = \psi_{II} - \psi_2$$

By introducing (14), (37) can be re-written as

$$\pi_1 - \pi_2 = \psi_{1.I} + \psi_{1.II} + \psi_{II.2} + \psi_{2.I} \quad (38)$$

where $\pi_1 - \pi_2$ is the electromotive force of the cell. Under certain conditions (which we shall suppose satisfied) $\psi_{1.II}$ can be taken as zero, just as in the case of two solutions of the same salt. If we also introduce (17), (38) becomes

$$\pi_1 - \pi_2 = (\psi_{1.I} + \psi_1^*) - (\psi_{2.II} + \psi_2^*). \quad (39)$$

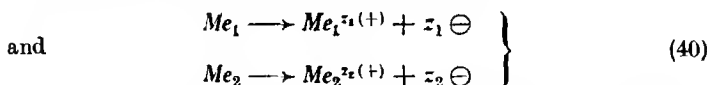
This equation gives an experimental means of relating two functions, each of which refers to one of the metals concerned. The potentials $\psi_{1.I}$ and $\psi_{2.II}$ will of course depend upon the ionic concentrations in I and II according to equation (32), but this effect can be eliminated by making the ionic activity the same in each case, *e.g.*, by making $a_1 = a_2 = 1$. In this case $\psi_{1.I}$ and $\psi_{2.II}$ will depend only on the metals and the solvent, and $-\psi_{1.I}$ and $-\psi_{2.II}$ can be considered as a measure of the tendency of the metals to pass into the ionic state in the solvent in question. As shown in equation (17), $-\psi_1^*$ and $-\psi_2^*$ represent the tendency of the metals to emit electrons to a second constant phase. The sums

$$-(\psi_{1.I} + \psi_1^*)$$

and

$$-(\psi_{2.II} + \psi_2^*)$$

thus represent the tendencies for the reactions



The right-hand side of (39) consists of the difference of these two sums: this is in agreement with the reaction taking place in the cell, which may be considered as the difference between the two electrode functions given in (40).

The left-hand side of (39) consists of the difference between the electromotive potentials π_1 and π_2 . According to (18) and (31) the complete expressions for these potentials are

$$\left. \begin{aligned} \pi_1 - \psi_I &= \psi_{1.I} + \psi_1^* \\ \pi_2 - \psi_{II} &= \psi_{2.II} + \psi_2^* \end{aligned} \right\} \quad (41)$$

The differences

$$\left. \begin{aligned} \pi_1' &= \pi_1 - \psi_I \\ \pi_2' &= \pi_2 - \psi_{II} \end{aligned} \right\}$$

when referred to unit ionic concentrations are commonly described as the *standard electrode potentials* of the metals in question. Since

it is impossible to measure the absolute value of a single potential difference, we are always concerned with the differences $\pi_1 - \pi_2$ or $\pi_1' - \pi_2'$. In all combinations measured in practice where the solutions involved are brought to the same potential [n., o.], π and π' can be used interchangeably in the equations for the potential.

If the standard potential of a given electrode is arbitrarily taken as zero, then by using this electrode in combination with others the standard electrode potentials of other metals may be obtained by using (39). If the metals are arranged in order of their standard electrode potentials the so-called *electrochemical series* is obtained. The standard potentials given in Table 6 are based on the assumption that the standard electrode potential of hydrogen is zero.

TABLE VI. *The Electrochemical Series*

Metal	Ion	Volts
<i>Na</i>	<i>Na</i> ⁺	2.7
<i>Mg</i>	<i>Mg</i> ⁺⁺	2.35
<i>Zn</i>	<i>Zn</i> ⁺⁺	0.76
<i>Fe</i>	<i>Fe</i> ⁺⁺	0.43
<i>Pb</i>	<i>Pb</i> ⁺⁺	0.12
<i>H</i>	<i>H</i> ⁺	0
<i>Cu</i>	<i>Cu</i> ⁺⁺	-0.34
<i>Ag</i>	<i>Ag</i> ⁺	-0.80
<i>Hg</i>	<i>Hg</i> ₂ ⁺⁺	-0.80
<i>Au</i>	<i>Au</i> ⁺	-1.4

The sign of the potentials given in this table is arbitrary, and can be chosen by convention. Remembering that $-(\psi_{1,1} + \psi_1^*)$ is a measure of the tendency of the metal to undergo reaction according to equation (40), it will be seen that if the solution receiving the metal ions is uncharged, then $-\pi_1$ will express the tendency of the metal to ionise according to (40). This tendency is great for the strongly "electro-positive" metals, and small for the less electro-positive ones. For this reason the electrode potentials in common use are obtained from the electromotive potentials occurring in the equations by a change of sign, so that the figures in the table represent $-\pi$. In other words, strongly electropositive metals are arbitrarily assigned a higher standard potential than weakly electro-positive ones. It should, however, be noted that other writers on electrochemistry quite often use the opposite convention as to signs.

The standard electrode potentials are generally regarded as a measure of the tendency of the metal atoms to go into solution as

positively charged ions. It will however be clear from the above treatment that it is not possible to compare ionisation tendencies in this way. The quantity which can be determined relatively is the *dissociation tendency* of the metal as exemplified by (40), where the metal ion goes over into a solvent (the chemical nature of which can vary and will affect the potential observed), while the electron is transferred to a fixed medium, which may be a vacuum. It is thus seen that it is a stoichiometric process taking place at the interface which determines and defines a single electromotive potential just as in the case when several such potentials are combined to form a galvanic cell.

The complete expression for the electrode potential (*i.e.*, the electromotive potential) of a metal-solution system is obtained by combining (32) with (41). The resulting expression can be written

$$\pi_1 - \psi_1 = \frac{RT}{z_1 F} \ln a_1 + \pi_{1(0)} - \psi_{1(0)},$$

or, if we can assume that $\psi_1 = \psi_{1(0)}$,

$$\pi = \frac{RT}{zF} \ln a + \pi_0 \quad (42)$$

This is identical in form with (33), but it will be seen that $-\pi_0$ is now a constant characteristic of the metal and equal to the standard electrode potential of the metal.

If the necessary conditions are satisfied, the ionic activity in (42) can be replaced by the ionic concentration. If this is not permissible, then the activity coefficients must be taken into account.

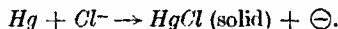
It will be seen from the above calculations that a tenfold change in the ionic activity at 20° causes a potential change of 0.0581 volts if the metal is univalent, and a change of half this magnitude if the metal is divalent.

A derivation of equation (42) for the electrode potential can also be obtained on the basis of a kinetic picture, the equilibrium between the metal electrode and the solution being considered as a mobile equilibrium in which the number of metal ions leaving the surface of the metal per unit time is the same as the number of ions condensing upon it (*cf.* the corresponding treatment of vaporisation equilibrium). The application of the Boltzmann distribution law to this phenomenon leads to equation (42). Since ions in solution are usually hydrated, it is important to note that this kinetic calculation can also be carried out on the assumption of hydrated metal ions, in which case the condensation of ions on the metal surface must be accompanied by the liberation of the bound solvent molecules. Equation (42) therefore does not require that the dissolved metal ions should actually be present in the free state.

(i) **Electrodes of the Second Kind.** This term is used to describe

a combination of a metal electrode with a solution which is in equilibrium with a sparingly soluble salt of the metal and which also contains an excess of the anion of this salt. An example is provided by an electrode of metallic mercury in contact with a potassium chloride solution containing mercurous chloride as a solid phase. This combination is commonly termed a "calomel electrode." The simple metallic electrodes described in the last section are correspondingly termed electrodes of the first kind.

In an electrode of the first kind the passage of current from metal to solution causes the solution to receive metal ions from the electrode (*e.g.*, $Hg \rightarrow Hg^+$). In an electrode of the second kind, on the other hand, the passage of current in the same direction causes the anion of the sparingly soluble salt to be removed from the solution by deposition of this salt, the process being, *e.g.*,



The potential of electrodes of the second kind must therefore depend on the anion concentration in the same way as the potential of electrodes of the first kind depends on the concentration of metal ion.

The expression for the potential can be easily derived in the following way. Equation (42) can of course be applied to both kinds of electrodes, and in the present case it becomes

$$\pi = \frac{RT}{F} \ln a_{Hg^+} + \pi_{0(+)} \quad (43)$$

Further, if the liquid is saturated with $HgCl$ we have

$$a_{Hg^+} a_{Cl^-} = \text{constant}.$$

Introducing this in (43), we obtain

$$\pi = - \frac{RT}{F} \ln a_{Cl^-} + \pi_{0(-)}, \quad (44)$$

where $\pi_{0(-)}$ is the electromotive potential of the electrode for $a_{Cl^-} = 1$. If $t = 20^\circ$ and a_{Cl^-} can be replaced by m_{Cl^-} , this becomes

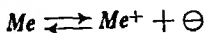
$$\pi = - 0.0581 \log m_{Cl^-} + \pi_{0(-)}. \quad (45)$$

It will easily be seen that in the general case when the valency of the ion is z (taken with the correct sign) the equation for the potential becomes

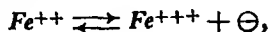
$$\pi = \frac{RT}{zF} \log a + \pi_{0(-)}, \quad (46)$$

in complete analogy with equation (39) for electrodes of the first kind.

(j) **Oxidation-reduction Electrodes.** In the cases treated above the tendency of the metals to dissolve as positive ions and the consequent equilibrium between the metal and its dissociation products



gives rise to an electric potential. Similarly, electric potentials may be produced in other equilibria involving electrons when the substances taking part are of a different charge type. For example, in a solution containing ferrous and ferric ions the equilibrium



will be set up, and in the presence of an "inert" metal electrode will give rise to a well-defined potential, since the metallic conductor may be regarded as a good solvent for electrons and will take them up. We may assume that the ferrous ion gives up an electron to the inert metallic electrode, being itself oxidised to a ferric ion, or that the reverse process takes place. In either case the electrode will be charged until the equilibrium potential is reached.

Electrodes of this type involving dissolved or non-metallic components are termed *oxidation-reduction* or "*redox*" electrodes. This terminology is related to the general oxidation-reduction scheme, which may be written



It is, however, clear that all primary electrode reactions (including the case of simple metallic electrodes dealt with above) can be represented by a scheme similar to the above, and might therefore be described by the same term.

The equation for the potential of an oxidation-reduction electrode can be derived by a method analogous to that used for a simple metal electrode. We can apply the general equilibrium condition VI. (8) to electrolytic systems by introducing the electrochemical potential in place of the chemical potential, giving

$$\sum \bar{\mu}_i = 0, \quad (47)$$

where the terms referring to reactants and resultants must of course have opposite signs in the summation.

Let the inert metallic electrode be phase (1) and the solution phase (I) while the chemical potentials are μ_{Red} , μ_{Ox} and μ_{\ominus} . The electrochemical potentials which determine the equilibrium are then respectively

$$\begin{aligned}\bar{\mu}_{Red(I)} &= \mu_{Red(I)} + z_{Red}F\psi_I \\ \bar{\mu}_{Ox(I)} &= \mu_{Ox(I)} + z_{Ox}F\psi_I \\ \bar{\mu}_{\ominus(I)} &= \mu_{\ominus(I)} - F\psi_I.\end{aligned}$$

Inserting these relations in the equilibrium condition

$$-\bar{\mu}_{Red(I)} + \bar{\mu}_{Ox(I)} + \bar{\mu}_{\ominus(I)} = 0 \quad (48)$$

we obtain

$$-(\mu_{Red(I)} + z_{Red}F\psi_I) + (\mu_{Ox(I)} + z_{Ox}F\psi_I) + (\mu_{\ominus(I)} - F\psi_I) = 0,$$

$$\text{or, since } z_{Ox} - z_{Red} = 1 \text{ and } \mu_{\ominus(I)} = -F\psi_1^* \quad .$$

$$\mu_{Ox(I)} - \mu_{Red(I)} + F\psi_I - F\psi_I - F\psi_1^* = 0. \quad (49)$$

The contact potential $\psi_{I1} = \psi_I - \psi_I$ is thus given by

$$F\psi_{I1} = \mu_{Ox(I)} - \mu_{Red(I)} - F\psi_1^*$$

or, introducing IV. (59),

$$\psi_{I1} = \frac{RT}{F} \ln \frac{a_{Ox}}{a_{Red}} - \psi_1^* - \frac{I}{F}. \quad (50)$$

The potential thus depends on the ratio of the concentrations of the oxidising and reducing agents, the nature of the redox system (expressed by I) and the inert metal electrode (expressed by ψ_1^*). Introducing (18), we find for the electrode potential

$$\pi_1 - \psi_1 = \frac{RT}{F} \ln \frac{a_{Ox}}{a_{Red}} - \frac{I}{F} \quad (51)$$

which no longer involves ψ_1^* . This shows that the nature of the "inert" electrode has no effect on the redox potential, as may also be shown by energy considerations.

If the redox electrode is combined with a standard electrode for which the electromotive potential is arbitrarily taken as zero, then if there is no contact potential between liquids the measured potential difference $\pi_1 - \pi_2$ can be termed the electromotive oxidation potential and can be written in the form

$$\pi = \frac{RT}{F} \ln \frac{a_{Ox}}{a_{Red}} + \pi_0, \quad (52)$$

where π_0 is the *standard oxidation potential* corresponding to $a_{Ox}=a_{Red}$. It is easily seen that in the general case when the difference in charge between the oxidising agent and the reducing agent is z positive charges, the equation becomes

$$\pi = \frac{RT}{zF} \ln \frac{a_{Ox}}{a_{Red}} + \pi_0 \quad (53)$$

analogous to equation (42) for the potential of a simple metal electrode.

The negative value of the oxidation potential is termed the *reduction potential*. Table VII gives the *standard reduction potentials* for a number of redox systems, the standard potential of hydrogen being taken as zero just as in Table VI. In agreement with the definition of the standard potential, the figures in general refer to equal concentrations of oxidising and reducing agent. In some cases, however, this convention is either inconvenient or impossible. Thus in Table VII, *Ag* and *I* refer to these elements in the solid state, *Br* to liquid bromine and *H₂* and *Cl₂* to gases at a pressure of one atmosphere. *Hy* and *Qu* refer to hydroquinone and quinone respectively, and *H⁺* represents a molar solution of hydrogen ions.

TABLE VII. *Reduction Potentials*

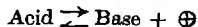
Red	Ox	Volts
<i>Ag</i> + 3 <i>CN</i> ⁻	<i>Ag(CN)₃</i> ⁻	0.51
<i>Cr</i> ⁺⁺	<i>Cr</i> ⁺⁺⁺	0.41
<i>H₂</i>	2 <i>H</i> ⁺	0
<i>Cu</i> ⁺	<i>Cu</i> ⁺⁺	-0.18
<i>Fe(CN)₆</i> ⁻⁻	<i>Fe(CN)₆</i> ⁻⁻⁻	-0.44
2 <i>I</i> ⁻	<i>I₂</i>	-0.58
<i>Hy</i>	<i>Qu</i> + 2 <i>H</i> ⁺	-0.70
<i>Fe</i> ⁺⁺	<i>Fe</i> ⁺⁺⁺	-0.75
2 <i>Br</i> ⁻	<i>Br₂</i>	-1.08
<i>Au</i> ⁺	<i>Au</i> ⁺⁺⁺	-1.2
<i>Tl</i> ⁺	<i>Tl</i> ⁺⁺⁺	-1.24
2 <i>Cl</i> ⁻	<i>Cl₂</i>	-1.35
<i>Ce</i> ⁺⁺⁺	<i>Ce</i> ⁺⁺⁺⁺	-1.6
2 <i>F</i> ⁻	<i>F₂</i>	-2.8

Just as in the case of the potential of a simple metallic electrode, it is clear that if the solutions are sufficiently dilute we can introduce concentrations in place of activities, while if this is not permissible the activity coefficients of the substances concerned must be introduced when calculating the potentials.

The scheme

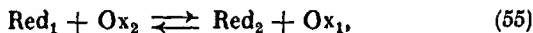


shows a complete analogy with the acid-base scheme



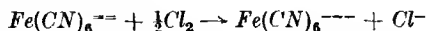
[VII. 3. b.], and it is possible to treat the two systems by exactly analogous methods. In practice, however, it is usual to employ electric potentials for characterising redox systems and equilibrium constants for characterising acid-base systems. This is chiefly due to the circumstance that water, the commonest solvent, possesses acidic and basic character and hence undergoes protolytic reactions with acids and bases, while it is not able to act as an oxidising or reducing agent to any measurable extent by losing or gaining electrons. In other words, water behaves towards redox systems in the same way as hexame (or other aprotic solvents) towards acid-base systems. Very powerful oxidising or reducing agents (e.g., fluorine or sodium) may be able to cause the water molecule to gain or lose an electron, but in this case further reactions take place with the evolution of oxygen and hydrogen respectively.

On uniting two corresponding redox systems an equilibrium is set up analogous to a double acid-base equilibrium,

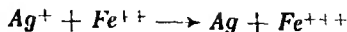


and it is possible to study such equilibria colorimetrically if one of the systems is a *redox indicator* (e.g., methylene blue) corresponding to an acid-base indicator.

A table of reduction potentials also provides the necessary information for predicting the direction of such double reactions. For example, we from the above numerical values that the ferrocyanide ion is a stronger reducing agent than the chloride ion, or, in other words, that chlorine is a stronger oxidising agent than the ferricyanide ion. Since the difference between the two potentials is about 0.9 volts, the process

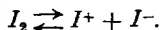


will go almost completely from left to right. On the other hand, the reaction



will lead to a state of measurable equilibrium, since Tables VI and VII show that the standard potentials have approximately the same value. The position of equilibrium can be calculated quantitatively from equation (52), since the values of π given by the two systems present in the mixture must of course be the same.

The electron transfer which is characteristic of redox equilibria can in some cases take place between identical atoms or molecules. This will lead to the production of ions and consequently electrolytic conductivity, just as in the case of autoprotolysis [VII. 3. b.]. For example, the electrical conductivity of iodine in inert solvents is attributed to the process



(k) **Gas Electrodes.** It has already been mentioned that a platinum electrode in contact with an atmosphere of hydrogen will function as a hydrogen electrode. For the electrode to act in a completely reversible manner it is necessary to "platinise" the platinum, i.e., to coat it with a layer of platinum black. Platinum and other noble metals will also function as reversible electrodes when brought into contact with other gases. Electrodes of this kind are termed *gas electrodes*, and do not differ in essence from the redox electrodes described above, so that the same equations are applicable.

For a hydrogen electrode (53) gives the expression

$$\pi_1 = \frac{RT}{F} \ln \frac{a_{H^+}}{p_{H_2}^{1/2}} + \pi_{1(0)},$$

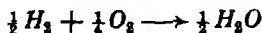
and for an oxygen electrode

$$\pi_2 = \frac{RT}{2F} \ln \frac{p_{O_2}^{1/2}}{a_{O^{2-}}} + \pi_{2(0)},$$

expressing the fact that a_{Red} for a hydrogen atom and a_{Ox} for an oxygen atom are proportional to the square root of the partial pressures. If the two electrodes are combined to give a cell, the electrode solutions being identical, the electromotive force of the cell is given by

$$\pi = \pi_2 - \pi_1 = \frac{RT}{F} \ln \frac{p_{H_2}^{1/2} p_{O_2}^{1/2}}{a_H + a_{O^{2-}}} + \pi_{2(0)} - \pi_{1(0)}. \quad (56)$$

According to the law of mass action the denominator of the logarithm has a constant value in any dilute aqueous solution, so that the electromotive force of the cell is independent of the nature of the solution. For example, the same value of π is obtained in both acid and alkaline solutions. This agrees with the fact that the electromotive process is the formation of water from hydrogen and oxygen, the passage of F units of electricity being associated with the process



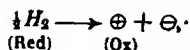
As would be expected from this reaction scheme and from equation (56), the electromotive force depends upon the pressures of the gases.

The oxygen electrode does not function very satisfactorily, but it has been found by indirect means that the electromotive force of the cell at 20° is $\pi = 1.23$ volts. This figure represents the affinity of formation of water [X. 1. a.]. Expressed in calories per gram-molecule of water, the value is

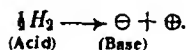
$$A = 2 \times 1.23 \times 23,060 = 56,700 \text{ calories,}$$

agreeing well with the value calculated thermodynamically by VI. (4) and VI. (37) from the dissociation of water vapour at high temperatures.

The hydrogen electrode reaction constitutes a special case, since it may be regarded either as a redox reaction according to the scheme



or as an acid-base reaction proceeding according to the identical scheme



This throws special light on the analogy between these two fundamental types of reaction. We may regard the proton as the simplest oxidising agent and the electron as the simplest base, while the hydrogen atom functions both as the simplest reducing agent and as the simplest acid.

(1) **Liquid-liquid Junction Potentials.** If a system consisting of two incompletely miscible liquids and a dissolved salt is at equilibrium, a potential will be set up at the interface between the two liquid layers. Just as in the case of a metal electrode, this potential will depend upon the ionic activities in the two phases. The relation can be derived from the fact that every ion taking part must have an identical electrochemical potential in the two phases.

We therefore write for the cation C and the anion A

$$\begin{aligned} \mu_{C(1)} + z_C F \psi_1 &= \mu_{C(2)} + z_C F \psi_2, \\ \mu_{A(1)} + z_A F \psi_1 &= \mu_{A(2)} + z_A F \psi_2, \end{aligned}$$

Introducing IV. (58) and writing $\psi_1 - \psi_2 = \psi$, this gives

$$\left. \begin{aligned} -RT \ln \frac{a_{C(1)}}{a_{C(2)}} &= z_C F \psi \\ -RT \ln \frac{a_{A(1)}}{a_{A(2)}} &= z_A F \psi \end{aligned} \right\} \quad (57)$$

where a is the absolute activity. It is of course impossible to determine ψ thermodynamically, just as for other phase equilibria.

For a salt of the type of NaCl , $z_C = 1$, $z_A = -1$, giving as the equilibrium condition

$$\frac{a_{C(1)}}{a_{C(2)}} = \frac{a_{A(2)}}{a_{A(1)}}, \quad (58)$$

or, introducing VII. (6), VII. (8) and VII. (9),

$$\frac{m_{\text{Salt}(1)}}{m_{\text{Salt}(2)}} = \frac{f_{\pm(2)}}{f_{\pm(1)}}, \quad (59)$$

where f is the absolute activity coefficient. (59) shows that the distribution of the salt is determined by the mean activity coefficients.

For a salt of the type of BaCl_2 , $z_C = 2$ and $z_A = -1$. In this case we obtain the equilibrium condition

$$\frac{a_{C(1)}}{a_{C(2)}} = \left(\frac{a_{A(2)}}{a_{A(1)}} \right)^2 \quad (60)$$

also equivalent to (59), which is in fact a general equation independent of the type of salt. (58) and (60) can also be written

$$a_{\text{Salt}(1)} = a_{\text{Salt}(2)},$$

thus reducing to the well-known equilibrium condition that the absolute activity of the salt must be the same in the two phases.

(m) Membrane Potentials and Electrolytic Membrane Equilibria (Donnan). If two solutions of the same salt in the same solvent but of different concentrations are separated by a membrane permeable to one of the ions but not to the other, then a potential difference will be set up between the two cells, since the ion to which the membrane is permeable will tend to pass from the more concentrated to the less concentrated solution, thus charging the latter with the appropriate sign. However, provided no other salts are present no detectable changes in concentration will be produced for reasons already described.

If the membrane cell contains an arbitrary number of salts having some ions which will pass through the membrane and some which will not, then we shall in general expect a membrane potential to be set up. The condition for electrochemical equilibrium is the identity of the electrochemical potentials. Hence when equilibrium is set up the following equation must be satisfied for each ion to which the membrane is permeable,

$$\mu_{1(1)} + \psi_1 z_1 F = \mu_{1(2)} + \psi_2 z_1 F,$$

where the index 1 in z_1 and μ_1 refers to the particular ion considered, and (1) and (2), ψ_1 and ψ_2 refer to the two solutions. Putting $\psi_1 - \psi_2 = \psi$, we obtain as the general expression for the membrane potential

$$\psi = \frac{1}{z_1 F} (\mu_{1(2)} - \mu_{1(1)}), \quad (61)$$

or, if the pressure is the same on the two sides of the membrane,

$$\psi = \frac{RT}{z_1 F} \ln \frac{a_{1(2)}}{a_{1(1)}} = \frac{RT}{z_2 F} \ln \frac{a_{2(2)}}{a_{2(1)}} = \dots \quad (62)$$

These equations are entirely analogous to those derived above for the liquid-liquid equilibrium. However, in the latter case it was impossible to compare the activity coefficients of ions in different media by thermodynamic methods. This is no longer the case for membrane equilibria, since the medium is the same on the two sides of the membrane. If the solutions are sufficiently dilute we may put $f = 1$ throughout, while for somewhat more concentrated aqueous solutions we can use the known expressions VII. (16) and VII. (18) for the activity coefficients. In both cases it will be possible to calculate the membrane potential and the distribution of the ions. For the same reason the membrane potential is *measurable* as the difference between the electromotive potentials of two standard electrodes (*e.g.*, calomel electrodes) in contact with the solutions on the two sides of the membrane, the junction potential being eliminated. The electromotive force of this combination may be directly equated to ψ .

The relation between the ionic activities expressed by equation (62) can also be derived easily from the fact that the activities of the *salts* in equilibrium must be the same in the two solutions. If one molecule of salt contains ν_1 ions of species 1 and ν_2 ions of species 2, this condition gives

$$a_{1(1)}^{\nu_1} a_{2(1)}^{\nu_2} = a_{1(2)}^{\nu_1} a_{2(2)}^{\nu_2}, \quad (63)$$

where

$$\frac{z_1}{z_2} = - \frac{\nu_2}{\nu_1}. \quad (64)$$

Equation (62) is only valid if the pressure is the same on the two sides of the membrane. This condition can always be fulfilled if non-permeating substances are present on both sides of the

membrane, *e.g.*, if the solvent is non-permeating. In this case it is possible to calculate the equilibrium state, given the amounts of non-permeating substances on each side of the membrane, and the total amounts of permeating substances. This calculation is of course simplest if all the activity coefficients can be taken as unity, in which case (62) can be written

$$\psi = \frac{RT}{z_1 F} \ln \frac{m_{1(2)}}{m_{1(1)}} = \frac{RT}{z_2 F} \ln \frac{m_{2(2)}}{m_{2(1)}} = \dots \quad (65)$$

whence

$$\left(\frac{m_{1(1)}}{m_{1(2)}} \right)^{\frac{1}{z_1}} = \left(\frac{m_{2(1)}}{m_{2(2)}} \right)^{\frac{1}{z_2}} = \dots \quad (66)$$

For example, let us consider the case in which solution (1) contains the non-permeating anion A_3 , while the anion A_2 and the cation K_1 both permeating, are also present. Let the volume of both solutions be 1 litre, the total amounts of $[K_1 A_2] = [K_1 A_3] = 0.1$, while $z_{K_1} = -z_{A_2} = -z_{A_3} = 1$. Equations (65) and (66) then give for the equilibrium state, $m_{1(1)} = 0.1333$, $m_{1(2)} = 0.0667$, $m_{2(1)} = 0.0333$, $m_{2(2)} = 0.0667$, where the indices refer to the corresponding ions: further at $t = 20^\circ$, $\psi = 0.0175$ volts.

If the concentration of the non-permeating salt is large compared with the concentrations of the other solutes, equation (66) shows that the permeating salts will be almost completely displaced from the solution containing the non-permeating ion.

The assumption that certain solutes can pass through the membrane while the solvent cannot is not generally satisfied. We will now consider the usual case in which the membrane is permeable to solvent. If the non-permeating substances are only present on one side of the membrane, then (by analogy with the simpler case treated in [V. 4. a.]) equilibrium cannot be attained when the two solutions are subjected to the same pressure. In this type of system consisting of two solutions separated by a semi-permeable membrane we will for convenience describe the solution containing the non-permeating substance or substances as the "inner liquid," and the solution containing only permeating substances as the "outer liquid." We can then say that a given inner liquid (*i.e.*, a solution in which the temperature, pressure, electrical potential and concentrations of both permeating and non-permeating molecules (including ions) are completely defined) will correspond to a completely defined outer liquid, *i.e.*, a solution which is in equilibrium with the inner liquid and which has completely defined values for the temperature, pressure, electrical potential and concentrations

of the permeating substances. If the inner liquid is referred to as (1) and the outer liquid as (2), we will define the membrane potential ψ as

$$\psi = \psi_1 - \psi_2, \quad (67)$$

while in agreement with the earlier treatment [V. 4.] the osmotic pressure is

$$P = p_1 - p_2. \quad (68)$$

The existence of this interdependence between the properties of the inner liquid and the outer liquid can be shown in the following way. If the given solution contains n permeating electrically neutral substances or electrically neutral combinations of ions, the equilibrium conditions required the identity of the same number of chemical potentials in the inner and outer liquids. In the outer liquid, however, there are only $n - 1$ independently variable concentrations for fixing the n chemical potentials thus defined, and it is therefore only possible to adjust all the chemical potentials to the required values if the pressure is introduced as the n th variable. The compositions and pressures of the solutions being thus fixed, the electrical potential is also completely determined.

The problem in hand is to determine the composition, pressure and electrical potential of the outer liquid corresponding to an inner liquid of given composition, pressure and electrical potential. The osmotic apparatus which serves to measure p is shown in Fig. 2. Let us assume that the inner liquid is a solution containing the medium K_0 , the cation K_1 and the anion A_2 , all three of which will pass through the membrane,

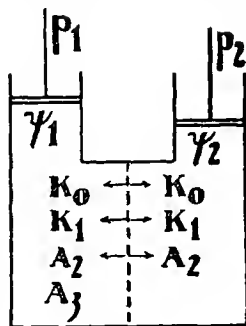


Fig. 2.

and in addition the anion A_3 , which will not pass through the membrane. Let the chemical potentials at zero pressure be $\mu_{0(1)}$, $\mu_{1(1)}$, $\mu_{2(1)}$, and $\mu_{3(1)}$ and the molecular volumes be $V_{0(1)}$, $V_{1(1)}$, $V_{2(1)}$ and $V_{3(1)}$, while p_1 and ψ_1 are the pressure and electric potential respectively. Analogous

terms with subscript (2) will be used for the outer liquid in equilibrium with the inner liquid (1). The effect of pressure on the chemical potential is obtained by integrating IV. (52), assuming the volume to be constant, giving

$$\left. \begin{aligned} \mu_{0(1)} + V_{0(1)}P_1 &= \mu_{0(2)} + V_{0(2)}P_2 \\ \mu_{1(1)} + V_{1(1)}P_1 + z_1F\psi_1 &= \mu_{1(2)} + V_{1(2)}P_2 + z_1F\psi_2 \\ \mu_{2(1)} + V_{2(1)}P_1 + z_2F\psi_1 &= \mu_{2(2)} + V_{2(2)}P_2 + z_2F\psi_2 \end{aligned} \right\} \quad (69)$$

Introducing the activity by IV. (59) and the osmotic pressure and membrane potential by (67) and (68), we have

$$\left. \begin{aligned} RT \ln \frac{a_{0(1)}}{a_{0(2)}} + V_0P &= 0 \\ RT \ln \frac{a_{1(1)}}{a_{1(2)}} + V_1P + z_1F\psi &= 0 \\ RT \ln \frac{a_{2(1)}}{a_{2(2)}} + V_2P + z_2F\psi &= 0 \end{aligned} \right\} \quad (70)$$

Since $a_{0(2)}$, $a_{1(2)}$ and $a_{2(2)}$ are the activities of the components of a given salt solution, they are not mutually independent, but are all determined by the composition of the salt solution. The three equations (70) therefore only contain three variables: P , ψ and the composition, so that the state of the system can be calculated completely from these equations. For the sake of simplicity we will assume that the solutions are ideal and that $z_1 = -z_2 = -z_3 = 1$. If x is the mole fraction, we then have

$$\begin{aligned} a_{1(1)} &= x_{1(1)} \\ a_{2(1)} &= x_{2(1)} \\ a_{3(1)} &= x_{3(1)} \\ a_{0(1)} &= 1 - (x_{1(1)} + x_{2(1)} + x_{3(1)}) \\ a_{1(2)} &= x_{(2)} \\ a_{2(2)} &= x_{(2)} \\ a_{0(2)} &= 1 - 2x_{(2)} \end{aligned}$$

In addition the condition for electrical neutrality gives $x_{1(1)} = x_{2(1)} + x_{3(1)}$. Introducing these relations and assuming that all the x 's are small compared with unity, we find

$$2RT[x_{(2)} - x_{1(1)}] + V_0P = 0 \quad (71)$$

$$RT \ln \frac{x_{1(1)}}{x_{(2)}} + V_1P + F\psi = 0 \quad (72)$$

$$RT \ln \frac{x_{2(1)}}{x_{(2)}} + V_2P - F\psi = 0 \quad (73)$$

These three equations are sufficient to determine the three unknowns, x_m , P and ψ . There is no simple explicit solution, but the equations can be solved numerically for any given case.

Equation (71) gives for the osmotic pressure,

$$P = \frac{2RT}{V_0}(x_{1(1)} - x_{2(2)}) = \frac{RT}{V_0}(x_{1(1)} + x_{2(1)} + x_{3(1)} - 2x_{2(2)}) \quad (74)$$

or, since $\frac{\Sigma x}{V_0} = m$ = the total volume molarity of the ions,

$$P = RT(m_{(1)} - m_{(2)}) \quad (75)$$

This equation expresses the general law that the osmotic pressure set up when two solutions are in direct osmotic contact is equal to the difference between the osmotic pressures of the solution measured against pure solvent.

Elimination of ψ between (72) and (73) gives

$$RT \ln \frac{x_{1(1)}x_{2(1)}}{x_{2(2)}^2} + (V_1 + V_2)P = 0, \quad (76)$$

so that if P is sufficiently small,

$$x_{1(1)}x_{2(1)} = x_{2(2)}^2 \quad (77)$$

agreeing with (66). This equation is thus equally valid whether the failure of the solvent to pass through the membrane is due to the impermeability of the latter or to the absence of any tendency for the solvent to diffuse. Under these conditions the membrane potential will also be given by the same equation (62).

If $x_{3(1)}$, the concentration of the non-permeating ion is large compared with the concentration of the permeating anion, i.e.,

$$x_{3(1)} \gg x_{2(1)},$$

equation (74) reduces to

$$P = \frac{2RT}{V_0}x_{3(1)} = 2RTm, \quad (78)$$

where m is the molarity of the non-permeating ion. We therefore obtain in this case the normal osmotic pressure of the binary electrolyte. If on the other hand

$$x_{2(1)} \gg x_{3(1)},$$

(77) gives

$$x_{2(2)} - x_{2(1)} = \frac{x_{3(1)}}{2}.$$

and hence from (74)

$$P = \frac{RT}{V_0} x_{s(1)} = RTm \quad (79)$$

since $x_{1(1)} = x_{s(1)} + x_{3(1)}$. The observed osmotic pressure can therefore be considered to arise entirely from the concentration of the non-permeating ion, since P is only half as great as in the previous case. This result is valid in general independent of the valency type: i.e., if the solvent is a salt solution having permeating ions at a concentration great compared with that of the non-permeating ion, then the osmotic pressure is given by equation (79). The proof of this statement will of course depend on the assumption made above that the activities in equation (70) can be replaced by concentrations: this assumption is, however, particularly allowable when a strong salt solution is used as solvent [VIII. 3. h.] as in the present case. A more general treatment of osmotic pressure in mixed solvents is given in [XI. 2. h.].

If $x_{2(1)}$ and $x_{s(1)}$ are of the same order of magnitude the osmotic pressure will be intermediate between the values obtained in the two limiting cases mentioned above.

If water or some other protolytically active solvent is used, the non-uniform distribution of ions in membrane cells of the types described above will usually lead to a difference in the acidity of the two solutions when equilibrium has been attained. For example, if compartment (1) contains an aqueous solution of a protolytically inactive salt with a non-permeating anion, while the second compartment (2) initially contains pure water, then when the permeating cation passes through the membrane it will carry hydroxyl ions with it so that the solution in (1) becomes acid relative to the solution in (2). The reverse will of course be the case if the cation is the non-permeating ion.

As examples of semi-permeable membranes we may mention parchment and nitro- or acetyl-cellulose. These membranes allow the passage of water and of ions or molecules of a moderate size, while large (colloidal) ions or molecules cannot pass through. For example, with the sodium salt of the dye-stuff Congo red the cation will pass through, but not the anion.

(n) **Diffusion Potentials.** In the treatment of single potentials so far given we have dealt entirely with equilibrium states. If two solutions of the same salt in the same solvent but of different concentrations are in direct contact we have a non-equilibrium system in which a potential may be set up by the tendency of the ions to become uniformly distributed throughout the whole system (Nernst). This potential is termed a *diffusion potential*.

The diffusion of the ions will take place in the direction of decreasing concentration. Since the mobility is different for different ions, this diffusion will lead to a partial separation, the faster ion gaining on the slower as they pass into the more dilute solution. This solution will therefore be charged with the same sign as the faster ion, while the more concentrated solution will be charged with the same sign as the slower ion. For example, if the mobility of the cation, u , is greater than that of the anion, v , then there will be a diffusion potential in which the dilute solution is positive to the more concentrated one.

The separation of the ions brought about by this effect is however much too small to be detected analytically, since (just as in the ordinary electrode equilibrium) even a very small displacement of the ionic concentrations will set up very large electrostatic forces opposing this displacement. The potential set up will retard the faster ion and accelerate the slower one. Since the net result is that the salt diffuses as a whole, the product of the mobility and the force acting on the ion in the diffusing system must be the same for each ion, i.e.,

$$K_k u = K_A v, \quad (80)$$

where K is the force.

This consideration can easily be used to obtain an expression for the potential difference. It must first be noted that the boundary between the two solutions is not sharp, but that there is a boundary layer of finite thickness in which the concentration changes continuously from c_1 to c_2 , while the potential changes from ψ_1 to ψ_2 and the distance from a given plane parallel to the layer from l_1 to l_2 . It has been previously shown [III. 2. c.] that in a gas of varying concentration the force acting on a molecule on account of the concentration difference is given by

$$K_1 = -kT \frac{d \ln c}{dl}, \quad (81)$$

where K_1 is the force in the direction of increasing l . The derivation of this equation shows that it can also be applied to a solution for which the gas laws are valid. In addition to this an ion will be subjected to a force on account of the electrical potential difference, given by

$$K_2 = -ze \frac{d\psi}{dl}, \quad (82)$$

where e is the charge on a univalent positive ion. For one gram-molecule of ions at a molarity m the total force is thus

$$K = -RT \frac{d \ln m}{dl} - zF \frac{d\psi}{dl}.$$

Introducing this expression into equation (80), if z_1 and z_2 are the valencies of the cation and the anion respectively, we have

$$\begin{aligned} u(RT d \ln m + z_1 F d\psi) &= v(RT d \ln m + z_2 F d\psi), \\ \text{or,} \quad d\psi &= -\frac{RT}{F} \frac{u-v}{uz_1 - vz_2} d \ln m, \end{aligned} \quad (83)$$

If the salt is uni-univalent, this becomes

$$d\psi = -\frac{RT}{F} \cdot \frac{u-v}{u+v} d \ln m. \quad (84)$$

The less-concentrated solution will thus be charged positively relatively to the more concentrated one if the mobility of the cation u is greater than the mobility of the anion v .

By integrating (84) we find for the potential in the direction (1) \rightarrow (2),

$$\psi = \psi_{2.1} = -\frac{RT}{F} \frac{u-v}{u+v} \ln \frac{m_1}{m_2}. \quad (85)$$

For a given electrolyte the diffusion potential is thus determined by the ratio of the concentrations of the two solutions. The effect of the electrolyte is expressed by the fraction $\frac{u-v}{u+v}$. It will be seen that this fraction (and hence the potential) becomes zero if the two ions have the same mobilities: this is very nearly true in some cases, *e.g.*, for the ions of potassium chloride. The maximum value of the fraction is unity, corresponding to the case in which the mobility of one ion is infinite compared with that of the other. The hydrogen ion is by far the fastest ion in aqueous solutions, and the diffusion potential may therefore be expected to be particularly large when acid solutions are used.

If a membrane is used to separate the two solutions of different concentrations the value of ψ in (85) will be altered on account of the changed mobility of the ions in the membrane. In the limiting cases $u = 0$ or $v = 0$ (85) reduces to expression (65) for the membrane potential.

If the gas laws are not applicable, the treatment given above will

lead to equations analogous to (84) and (85), but containing the activities of the ions in place of their concentrations.

Diffusion potentials can be measured in the same way as membrane potentials, i.e., by bringing the two solutions between which the potential is set up into contact with the liquid phase surrounding two metallic electrodes, the contact potential being eliminated. The diffusion potential is then equal to the difference between the electromotive potentials of the metal electrodes. The problem really consists of bringing two electrolyte solutions into contact in such a way that there is no potential difference between them. The same problem arises in the methods described above for measuring membrane potentials, redox potentials and comparing metal electrodes.

The diffusion potential depends on the difference in the mobilities of the two ions, and will disappear if u and v are equal. The insertion of a concentrated salt solution with ions of equal mobilities between two solutions may therefore be expected to eliminate the diffusion potential to a great extent. In practice a saturated solution of potassium chloride is used for this purpose, since u and v are almost equal for this salt (cf. Table I). The methods described above for measuring potentials are hence only accurate in so far as the precautions used for eliminating the diffusion potential are effective.

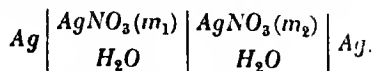
If two solutions of the same salt are in mutual contact and are each in contact with a metal electrode corresponding to the cation, the combination is termed a "concentration cell with transference." The potential difference between the electrodes is equal to the sum of the potential differences in the cell, and since the electrodes are of the same metal this is equal to the difference between the electromotive potentials of the electrodes. If both ions are univalent, (32) and (85) give for the present case

$$\pi = \pi_2 - \pi_1 = -\frac{RT}{F} \ln m_1 + \frac{RT}{F} \frac{u-v}{u+v} \ln \frac{m_1}{m_2} + \frac{RT}{F} \ln m_2,$$

or,

$$\pi = -\frac{2v}{u+v} \frac{RT}{F} \ln \frac{m_1}{m_2}. \quad (86)$$

This equation may, for example, be applied to the cell



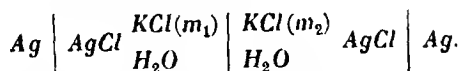
Equation (86) can also be derived by a purely thermodynamic argument, in the same way in which (25) was derived for a concen-

tration cell without transport. According to [1. c.], the electromotive process in the present case is the transfer of $\frac{v}{u+v}$ gram-molecules of salt from one electrode compartment to the other for the passage of F units of electricity through the cell.

If a corresponding cell is constructed from two electrodes of the second kind, the expression for the electromotive force of the cell becomes

$$\pi = \frac{2u}{u+v} \frac{RT}{F} \ln \frac{m_1}{m_2} \quad (87)$$

applicable, for example, to the cell



All these equations involving the concentration are, of course, only applicable for high dilutions of the ions concerned. If in place of water a *concentrated salt solution* is used as the solvent, then the activity and the concentration will be proportional and the simple gas laws will be applicable [VII. 2. c.]. In addition to this the diffusion potential will practically disappear, since at the point where the two liquids come into contact the conduction of electricity will be almost entirely due to the ions which are present in large excess and have the same concentration in both solutions. In this case (86) and (87) can be replaced by the simple and exact equation

$$\pi = \frac{RT}{zF} \ln \frac{m_1}{m_2}. \quad (88)$$

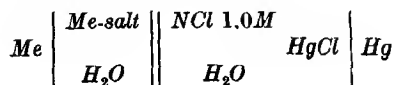
(o) **The Measurement of Single Potentials.** It has already been stated and illustrated by numerous examples that a "single potential," i.e., the electrical potential difference between two phases, can only be measured thermodynamically if the two phases are composed of the same "medium." Even under these conditions the potential can only be obtained with a certain degree of approximation, since even if the solvent (i.e., the predominant component) remains unchanged, any change in the composition involves a change in the phase as a whole, and hence a certain "medium effect." In dilute solutions, however, this effect can be neglected: thus the membrane potential has been mentioned above as a potential which can be measured accurately.

If a potential is set up between two phases in different media,

e.g., between a metal and a solution of a corresponding metallic salt, then no absolute measurement of that single potential is possible. It is however possible to carry out a comparative measurement by measuring the *difference* between two such potentials, provided that the solutions of the salts are dilute and are in the same solvent, and that these two solutions can be brought into contact without potential difference. This last condition is best fulfilled by using a *concentrated solution of another salt* as solvent throughout, the salts corresponding to the metal electrodes being added in concentrations which are small compared with that of the solvent salt. In a combination of this kind there will be no potentials other than those at the metal electrodes, and the effect of concentration or of the nature of the metal can be accurately studied. For example, the theoretical concentration variation expressed by equation (36) is found to apply with high accuracy. If it is not desired to work in such a salt solution, the liquid-liquid junction potential may be largely eliminated as described above by using a concentrated *KCl* solution as "bridge solution."

As a *standard electrode* for such measurements we may use a hydrogen electrode, or some other constant and reproducible electrode having a known potential relative to the hydrogen electrode. A frequent choice is the "calomel electrode," consisting of mercury covered by a layer of solid calomel, *HgCl*, in contact with molar or 0.1 molar *KCl* solution. If molar *KCl* solution is used, the electromotive potential of this solution is 0.281 volts greater than that of the standard hydrogen electrode.

For the relative determination of single potentials we may combine a series of "half-cells," *i.e.*, metal electrodes with their appropriate liquid phases, with a calomel electrode as the other half-cell, thus forming the combination,



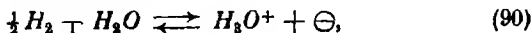
where the double line indicates that the junction potential has been eliminated. If *Me* = hydrogen, the electromotive force of the combination will be 0.281 volts with mercury as the positive pole.

(p) **The Electrometric Determination of Hydrogen Ion Concentrations.** Of the various methods which can be used for determining hydrogen ion concentrations, the electrometric method is of particular importance. The hydrogen ion concentration is calculated by means of equation (42) for the electrode potential, the estimation of the single potentials involved being carried out by the method

described under (o). In the present instance $z = 1$, and for $t = 20^\circ$ (42) can be written

$$\pi = 0.0581 \log a_{H^+} + \pi_0. \quad (89)$$

The nature of the hydrogen electrode (*cf.* (k)) shows that the electromotive force is a direct measure of the proton activity in the solution. Since in aqueous solution the hydrogen ion is H_3O^+ , the stoichiometric electrode reaction is



showing that the potential depends not only on the hydrogen pressure and the hydroxonium ion activity, but also on the activity of the solvent. The last factor is, however, unimportant in dilute solutions, since the activity of the solvent is practically constant.

The effect of the hydrogen pressure is however of importance in practice on account of variations in the barometric pressure. The magnitude of the effect can be calculated from equation (27), which gives for $t = 20^\circ$ and small pressure changes,

$$-\Delta\pi = 0.0126 \frac{\Delta p}{p} \quad (91)$$

showing that an increase of pressure by 8% corresponds to an increase of 1 millivolt in the electrode potential.

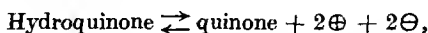
The potential of the hydrogen electrode is equally well defined in alkaline and in acid solutions. In alkaline solutions this potential is much higher, in agreement with the small concentration of hydrogen ions present. The difference between the potentials of the hydrogen electrode in 0.1 *NaOH* and 0.1 *HCl* at 25° amounts to 0.696 volts, whence equation (42) gives $m_{H_3O^+} = 1.70 \times 10^{-13}$ in the first solution, so that

$$K_{(H_2O)} = 1.70 \cdot 10^{-14}.$$

This value is considerably greater than the value 1.02×10^{-14} previously given [VII. 3. b.] for pure water at 25° . This is because the present value refers to a 0.1 molar electrolyte solution in which the dissociation constant is increased by the factor $\left(\frac{1}{f_1}\right)^2$ [VII. 3. h.].

Equation (90) represents the simplest form of an oxidation-reduction equilibrium in aqueous solution involving the proton. It will, however, be easily seen that any such oxidation-reduction reaction in which hydrogen ions take part will be affected by the

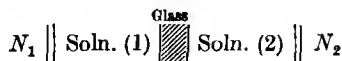
acidity, and can be used for determining acidity if the potential is set up sufficiently rapidly. An example of this kind of electrode which has considerable practical importance is the *quinhydrone electrode*, consisting of an electrode of platinum or some other inert substance surrounded by a solution of quinhydrone. The equilibrium scheme is then



showing that the quinhydrone concentration has no effect upon the potential, which depends only on the ratio between the concentrations of hydroquinone and quinone. The electrode will not, however, function at p_{H} values greater than about 8.5. As previously mentioned, the standard reduction potential of the quinhydrone electrode is 0.70 volts lower than that of the hydrogen electrode.

Another electrode which is used for determinations of hydrogen ion concentrations in aqueous solution is the so-called *glass electrode*, which has considerable advantages on account of its chemical inertness towards the solutions used. The electrode consists of a very thin-walled glass bulb filled with a solution (1) of well-defined acidity. This "inner liquid" is connected to a standard electrode or some other electrode of constant potential. The bulb is immersed in the "outer liquid" (i.e., the solution whose acidity is required), which is connected with a standard electrode of the same kind as the first. It is then found that (in the absence of diffusion potentials) the electromotive force of the whole combination varies with the hydrogen ion activity of the outer liquid according to the usual logarithmic formula (89).

The cell is represented by the scheme



where N_1 and N_2 represent the standard electrodes connected respectively to the inner liquid (1) and the outer liquid (2), the junction potentials being eliminated. The working of the cell is easily understood if it is assumed that the glass acts as a membrane permeable to hydrogen ions but not to any other type of ion. Equation (62) for the membrane potential then shows that the total electromotive force of the combination is given by

$$\pi_1 - \pi_2 = \psi_{(1)} - \psi_{(2)} = \frac{RT}{F} \ln \frac{a_{(2)}}{a_{(1)}}.$$

Since $a_{(1)}$ is constant, this reduces to the usual relation between π_2 and $a_{(2)}$, the hydrogen ion activity.

The resistance of the glass bulb is very high, and an electrostatic instrument must therefore be used for measuring the potential.

If ions other than the hydrogen ion were able to pass through the glass there would not in general be electrochemical equilibrium between the two solutions and the principle of the method would fail. There are no free protons or hydrogen ions present in the glass [VII. 3. b], and the fact that this ion in particular can penetrate is probably connected with the ampholytic nature of glass.

(q) **Polarisation and Over-voltage.** If an electric current passes through an electrolytic conductor a potential is usually set up between the electrodes so as to oppose the passage of the current. This phenomenon is termed *electrolytic polarisation*: it is a simple consequence of the chemical energy produced by the current, and must be equivalent to the corresponding loss in electrical energy. In certain cases the passage of the current may be said to produce two half cells, and if the current density is sufficiently low the polarisation potential can be calculated as the sum of the electromotive potentials of these two half-cells. Thus if zinc bromide is electrolysed slowly we find a polarisation of about 1.8 volts, equal to the difference of the standard potentials of Zn and Br. If, as in this case, the process consists of a simple decomposition, the polarisation is termed the *decomposition potential* of the salt, and the electrolysis is reversible.

The polarisation is of course dependent on the concentration of the salt being electrolysed and also the concentration of the decomposition products. In the case given above the initial bromine concentration round the anode will be zero, so that to begin with the polarisation at this electrode should theoretically be zero. As increasing amounts of bromine are liberated the polarisation increases, and the decomposition potential is reached at saturation concentration. For this reason the current passing through an electrolysis apparatus will not in general be a simple function of the applied potential. If the latter is appreciably smaller than the decomposition potential the current will be very small, and (as appears from the equation for the electrode potential) will vary rapidly with the potential. Once the decomposition potential has been reached the polarisation will be approximately constant, and the current will therefore be approximately proportional to the amount by which the applied potential exceeds the decomposition potential. If the electrolyte contains several substances with different decomposition potentials, this will appear on the current-potential curve as quite sharp changes in the ratio between the current and the potential. These differences may be used for separating substances with considerably different electrode potentials, *e.g.*, copper from zinc or iodine from chlorine.

If a salt solution is electrolysed between electrodes composed of the metal contained in the salt (*e.g.*, a silver nitrate solution between silver electrodes) an infinitesimal current needs only an infinitesimal potential, since electrolysis results only in the transfer of the metal from anode to

cathode. For larger potentials and currents, however, polarisation will be set up on account of the concentration change due to electrolytic transport, the concentration of salt round the anode being increased, and the concentration round the cathode decreased. The magnitude of this polarisation will depend on the relation between the current density and the velocity with which the salt diffuses. After a certain time a stationary state will be reached depending upon the conditions of experiment, in which the concentration changes caused by the current are compensated by those caused by diffusion of the salt. In this stationary state the nitrate ions will clearly be stationary, the current being entirely carried by the metal ions.

Special importance attaches to the irreversible polarisation phenomena associated with the electrolytic discharge of gases, *e.g.*, hydrogen and oxygen. For example, if we measure the potential difference between a reversible hydrogen electrode carrying no current and an inert electrode at which hydrogen is being liberated electrolytically, a difference will be found which depends on the current density and upon the nature of the inert electrode. It is always found that the reversible electrode exhibits a higher electromotive potential than the electrode at which hydrogen is being discharged, and this difference is termed the hydrogen *over-voltage*. Its magnitude for low current densities is about 0.44 volts for a mercury electrode, but is practically zero for an electrode of platinised platinum.

The existence of an over-voltage probably has no direct connection with the gaseous state of the substances taking part, but is rather due to the fact that the reversible electrode potentials of these gases refer to the diatomic molecules of the gas, while the ion is monatomic. The electrode process therefore consists not only in the simple discharge of an ion, but also in the combination of the liberated atoms to molecules. It is easily understood that this process takes place more readily on a platinum electrode than on electrodes of other metals, since it is known that hydrogen dissolves in platinum partly in the atomic state, so that in this metal a reversible equilibrium must be set up between hydrogen atoms and hydrogen molecules.

In the electrolytic preparation of *deuterium* (the isotope of hydrogen with mass 2) by the electrolysis of water, the separation from ordinary hydrogen is due to the fact that the heavier isotope has the greater over-voltage.

CHAPTER X

AFFINITY

THIS chapter deals partly with the *concept* of affinity, as it has been developed on the basis of observations of chemical processes, and partly with affinity expressed quantitatively in terms of the energy changes associated with processes involving molecules and atoms. These two concepts have been developed along very different lines, but they are sufficiently related to enable them to be treated under the same heading.

1. REACTION AFFINITY

(a) The Concept of Affinity. From early times the word "affinity" has been used to denote the chemical attraction between substances. This does not provide a satisfactory definition of the concept, but by considering certain reactions it was possible to arrange substances in order of increasing or decreasing affinity, so that if the affinity of *A* for *C* is greater than that of *B* for *C*, *A* will be able to replace *B* in the compound *BC* (Bergmann). It was later realised that the course of a reaction depends not only on the chemical nature of the substances, but also on the concentrations, thus leading to the laws of homogeneous equilibrium (Berthollet, Guldberg and Waage).

The first unequivocal definition of affinity is due to Julius Thomsen, who identified the affinity of a process with the heat evolved. According to this theory (often referred to as Berthelot's principle) every spontaneous process should be exothermic, and the amount of heat evolved should be a measure of the affinity. Actually, however, endothermic processes can also take place spontaneously, particularly at high temperatures, so that this theory cannot be maintained although it agrees with the facts for processes associated with *large* heat evolution. The theory is also definitely contradicted by thermodynamic results.

The fact that every spontaneous process can be made to perform work is really implied in the term "spontaneous," since this may be taken to mean a process which can be carried out without supplying work. Since further any chemical process by definition proceeds in the direction determined by affinity, the affinity and the

maximum work obtainable at a given temperature always have the same sign: hence it is reasonable to use the maximum work as a *quantitative* measure of the affinity (van't Hoff).

It has previously been pointed out that the decrease in free energy associated with a chemical reaction represents the maximum work obtainable from the reaction at the temperature in question. The concept of affinity is thus closely related to the thermodynamic functions. Since it is the work done by the *chemical* forces which we wish to express in the affinity, in determining it we do not take into account the work done by a volume change against a positive or negative pressure, which appears as the term $p dv$ in equation I. (31). (In the same way we exclude potential thermal energy from consideration by the condition that the temperature must be constant). We can therefore define the affinity as the change in free energy when the reaction in question takes place at constant temperature and constant volume. In agreement with I. (31) and I. (44) the affinity of the reaction $A \rightarrow B$ can be written

$$A^* d\alpha = -(dF)_{T,v} = -\sum \mu_1 dn_1, \quad (1)$$

where A^* is the affinity (*i.e.*, the chemical work, as defined above, for unit quantity of substance), α is the extent of the reaction, and dn is positive or negative for the A -system or the B -system respectively. The affinity can of course also be expressed in terms of the G -function, since on combining (1) with I. (44) we have

$$A^* d\alpha = -(dG)_{T,p} = -\sum \mu_1 dn_1. \quad (2)$$

In a chemical reaction the changes in the numbers of molecules are not mutually independent, since the quantities of the substances change in the proportions determined by the stoichiometric reaction equation. Equation VI. (7) gives

$$\left. \begin{aligned} dn_{A_1} &= -n_{A_1} d\alpha, \\ dn_{B_1} &= n_{B_1} d\alpha, \end{aligned} \right\} \quad (3)$$

where n_{A_1} , n_{A_2} , ..., n_{B_1} , n_{B_2} , ... are the stoichiometric ratios occurring in the equation. Introducing this relation, equations (1) and (2) become

$$A^* = -\left(\frac{\partial F}{\partial \alpha}\right)_{T,v} = -\left(\frac{\partial G}{\partial \alpha}\right)_{T,p} = \sum \mu_{A_1} n_{A_1} - \sum \mu_{B_1} n_{B_1} \quad (4)$$

As previously shown [VI. 2. a.], the standard states for chemical potentials cannot be chosen arbitrarily, but must be fixed in such a way that equations (1) to (4) are satisfied in an arbitrarily chosen state: in particular so that A^* is zero in an equilibrium state. We

thus see the fundamental importance of the quantity A^* in chemical processes, since these equations show that it may be regarded as a function measuring the tendency for the reaction or the maximum irreversibility of the process.

If the reaction is heterogeneous, or if for any other reason the chemical potentials are independent of α , then the affinity will not change as the process progresses. In that case the reaction goes to completion, and A^* can be simply put equal to the decrease in the G -function per unit quantity of substance when the process takes place at constant pressure and temperature. In homogeneous systems, on the other hand, the chemical potentials will change

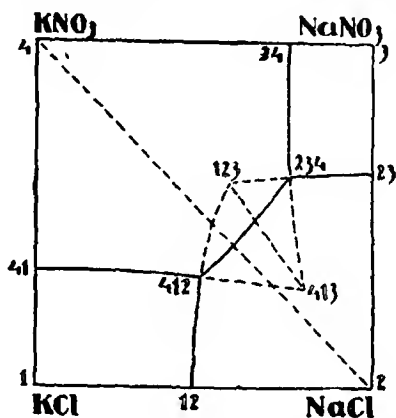


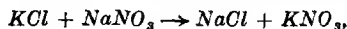
Fig 1.

with the composition of the system and the affinity will thus decrease during the process, finally becoming zero when equilibrium is reached.

(b) **The Principle of Affinity Determinations.** The affinity can be measured by any of the methods previously described for measuring the maximum work of a chemical reaction. The work corresponding to the term $p dv$ must then be subtracted, though for transformations involving solid and liquid substances this term is usually inappreciable. Special mention may be made of electrometric methods [IX. 2. c.-e.], and measurements depending on dissociation and equilibrium constants. In simpler cases, such as the affinity of changes of state or allotropic transformations, measurements of vapour pressure [II. 4. e.] or solubility [V. 3. a.] may be employed. The affinity of crystallisation of a supercooled liquid can also be determined by considering a reversible volume change until the equilibrium pressure is reached, followed by transformation at the

constant equilibrium pressure and subsequent reversible volume change to the initial pressure. The affinity is then given by the total mechanical work obtained in these operations, and can be calculated if the corresponding volumes and compressibilities are known.

As an example of the application of a special method (which can however be applied in principle in many cases) we shall describe the measurement of affinity in the transformation of a *reciprocal salt pair*, e.g.,



all substances being present in the crystalline state. Processes of this kind have been previously described (IV. 5. a.), and it was particularly emphasised that only one of the salt pairs in such a transformation will normally be stable.

To illustrate the method we will consider the diagram in Fig. 1, which shows schematically the solubility relationships in such a system. The axes of abscissæ and ordinates represent respectively the mole fractions of the cation and the anion in the solution, the amount of solvent not being taken into account in the calculation: thus, for example,

$$x = \frac{n'}{n' + n''},$$

$$y = \frac{n'''}{n''' + n''''},$$

where the total amounts of K , Na , Cl and NO_3 are respectively n' , n'' , n''' and n'''' . The four vertices of the diagram thus correspond to solutions of the four pure salts KCl , $NaCl$, $NaNO_3$ and KNO_3 (represented in the order given by the numerals 1, 2, 3 and 4) which of course only contain two ions. Similarly, the sides of the square correspond to the presence of two salts containing three ions, and point inside the square to solutions containing all four ions. In each case the relative quantities of the ions present are determined by the position of the appropriate point in the diagram.

The curves in the diagram represent solubility curves corresponding to simultaneous saturation with two salts, while the points of intersection of these curves correspond to simultaneous saturation with three salts. Stable states are represented by full curves and unstable states by dotted curves. The solid phases corresponding to the curves and their points of intersection are shown on the figure.

The stable 2-salt curves intersect at the stable 3-salt points 4.1.2 and 2.3.4, while their unstable continuations intersect at the unstable 3-salt points 1.2.3 and 4.1.3. The system 2.4 (i.e., $NaCl + KNO_3$) is thus the stable salt pair, while the system 1.3 is unstable: this is also indicated by the direction of the arrow in the reaction scheme. If a solvent is present to catalyse the process, the salts $KCl + NaNO_3$ will be transformed into the reciprocal system.

This process can be carried out reversibly by transferring the salt 1 from the point 4.1.2 to the point 2.3.4, i.e., from the first to the second of the corresponding saturated solutions. These solutions are composed

of water, the salts 2 and 4 with which the solutions are saturated, and salt 1: at 4.1.2 the solution is saturated with this last salt, while at 2.3.4 (and at any point in the triangle 2.3.4) it may be said to be present in *negative amount*.

We can now apply the Gibbs-Duhem equation [IV. 3. c.]

$$m_1 d\mu_1 + m_2 d\mu_2 + m_4 d\mu_4 + m_{H_2O} d\mu_{H_2O} = 0$$

to these solutions. On account of the saturation equilibrium along the curve 4.1.2—2.3.4, we have

$$d\mu_2 = d\mu_4 = 0,$$

giving the relation

$$m_1 d\mu_1 + m_{H_2O} d\mu_{H_2O} = 0,$$

or, introducing

$$x = \frac{m_1}{m_1 + m_{H_2O}}$$

$$x d\mu_1 + (1 - x) d\mu_{H_2O} = 0,$$

and hence

$$\mu_1 = - \int_{x(4.1.2)}^{x(2.3.4)} \frac{1-x}{x} d\mu_{H_2O}.$$

In order to determine the potential difference $\mu_{1(4.1.2)} - \mu_{1(2.3.4)}$, the above expression must be integrated, which is possible if the vapour pressure of the water is known as a function of x . The affinity can thus be found if vapour pressure measurements are available along the whole curve 4.1.2—2.3.4, since

$$A^* = \mu_{1(4.1.2)} - \mu_{1(2.3.4)}$$

(c) **The Variation of Affinity with Temperature.** To determine the temperature variation of the affinity of the process $A \rightarrow B$ we shall apply equation I. (37) to the two systems concerned and then subtract the two equations obtained. We thus find

$$\Delta G = \Delta H + T \left(\frac{d\Delta G}{dT} \right)_p, \quad (5)$$

where $\Delta G = G_B - G_A$, and similarly for ΔH . If the reaction is of the heterogeneous type dealt with above, the affinity is equal to $-\Delta G$. Further, ΔH is in general equal to Q , the heat absorbed in the process, so that (5) can be written

$$A^* = -Q + T \frac{dA^*}{dT}. \quad (6)$$

It is thus seen that the relation between the affinity and the heat change is not as simple as assumed in the Thomsen-Berthelot principle, which demanded $A^* = -Q$, since (6) shows that this

identity is only true at absolute zero and at temperatures for which $dA^*/dT = 0$. If A^* increases with increasing temperature the affinity is greater than the heat evolved, while if it decreases the reverse is the case. By differentiating (6) with respect to temperature we obtain

$$\frac{dQ}{dT} = T \frac{d^2 A^*}{dT^2} \quad (7)$$

showing that the curvature of the A^* -curve is determined by the sign of the temperature coefficient of the heat evolution, so that

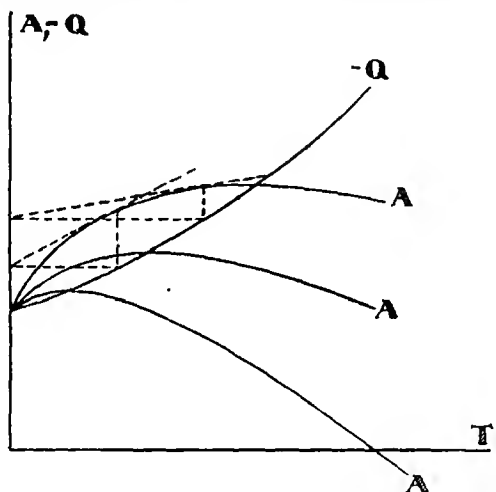


Fig. 2

when this temperature coefficient is zero the A^* -curve will have a point of inflexion.

If the variation of Q with temperature is known right down to absolute zero it is possible to construct a series of A^* -curves which satisfy equations (6) and (7). Such curves are shown in Fig. 2, which also illustrates the geometrical construction for constructing the $-Q$ -curve from a given A^* -curve on the basis of I. (37). It appears from this construction, and also from I. (39), which can be written in the form

$$A^* = T \int \frac{Q}{T^2} dT + i,$$

that A^* and Q coincide at $T^* = 0$, but that there are an infinite number of A^* -curves consistent with a given Q -curve. The problem

of determining the affinity of a chemical process from a knowledge of the Q -curve (*i.e.*, from thermal measurements alone) is therefore one which cannot be solved by classical thermodynamics.

However, in certain cases, namely reactions involving only pure substances in the *crystalline state*, it becomes possible to determine the affinity by introducing a special assumption about thermodynamic behaviour at low temperatures. Since this assumption is of a very general nature and has been verified over a wide field, it is termed the third law of thermodynamics.

(d) **The Third Law of Thermodynamics.** According to equation I. (30),

$$G = H - TS. \quad (8)$$

Applying this equation to the reaction $A \rightarrow B$, putting $G_B - G_A = \Delta G$, etc., we obtain

$$\Delta G = \Delta H - T\Delta S,$$

or, since $A^* = -\Delta G$ and $Q = \Delta H$,

$$A^* = -Q + T\Delta S. \quad (9)$$

The affinity can thus be determined from the heat evolved in the process provided that ΔS for the process can also be measured. Since $T\Delta S$ is equal to the heat absorbed when the process takes place reversibly, the affinity can only be determined by this method when the process can be carried out reversibly. This is not always possible in practice, and in cases where the reversible process can be realised it is in general much easier to measure the work done (*i.e.*, A^*) than the heat absorbed. The use of equation (9) is hence really equivalent to a direct measurement of the affinity. Further, a determination of ΔS in this way does not really constitute a solution of the problem outlined above, which was to determine the affinity from *thermal* data alone. If therefore equation (9) is to be used to determine the affinity, ΔS must be found by some other means.

For this purpose we write ΔS in the form

$$\Delta S = S_B - S_{B_0} - (S_A - S_{A_0}) + \Delta S_0, \quad (10)$$

where the subscript 0 refers to absolute zero. The entropy differences referring to the two separate systems are directly accessible from *calorimetric* measurements, since by using I. (62),

$$T \left(\frac{\partial S}{\partial T} \right)_p = c_p$$

we find

$$S - S_0 = \int_0^T c_p d \ln T. \quad (11)$$

Equation (10) can therefore be re-written as

$$\Delta S = \int_0^T \Delta c_p d \ln T + \Delta S_0, \quad (12)$$

where $\Delta c_p = c_A - c_B$ is the increase in heat capacity associated with the process in question.

The contribution of the third law to the problem of determining ΔS in (9) (and hence the affinity) lies in the postulate

$$\Delta S_0 = 0, \quad (13)$$

i.e., it is assumed that a reaction between pure crystalline substances at absolute zero takes place *without entropy change*. Equation I. (34) shows that this postulate can also be written in the form

$$-\lim_{T \rightarrow 0} \left(\frac{d\Delta G}{dT} \right) = \lim_{T \rightarrow 0} \left(\frac{dA^*}{dT} \right) = 0. \quad (14)$$

The content of these equations is known as *Nernst's heat theorem*. Both expressions can be derived from the more drastic assumption that the absolute value of the entropy of any pure crystalline substance is zero at absolute zero (Planck).

Since the limiting value of dA^*/dT is zero, the value of A^* in the vicinity of $T = 0$ can be written in the form

$$A^* = A_0^* + \alpha T^\beta, \quad (15)$$

where $\beta > 1$. Introducing this equation in (7), we have

$$\frac{dQ}{dT} = \alpha \beta (\beta - 1) T^{\beta-2},$$

an expression which becomes equal to zero when $T = 0$ provided that $\beta > 1$. We thus see that the heat theorem predicts that Q will behave in the same way as A^* , i.e.,

$$\lim_{T \rightarrow 0} \left(\frac{dQ}{dT} \right) = c_{B_0} - c_{A_0} = 0,$$

where c_{A_0} and c_{B_0} are the specific heats of the two systems at absolute zero. This is in complete agreement with the quantum theory of the specific heats of crystals, and also with experimental data. Equation II. (43) states that at very low temperatures the

specific heat is proportional to the third power of the absolute temperature. This leads to

$$Q = Q_0 + 3\alpha T^4. \quad (16)$$

(where α may have either positive or negative values), and hence for the same range of low temperatures

$$A^* = A_0^* + \alpha T^4. \quad (17)$$

The complete A^* -curve can be derived from the heat theorem by putting $\Delta S_0 = 0$ in (12), giving

$$\Delta S = \int_0^T \Delta c_p d \ln T \quad (18)$$

which expression can be substituted in (9). This gives

$$A^* = -Q + T \int_0^T \Delta c_p d \ln T, \quad (19)$$

while according to I. (60) the curve for the heat evolution can be written

$$Q = Q_0 + \int_0^T \Delta c_p dT. \quad (20)$$

Figs. 3 and 4 show the course of the A^* - and $-Q$ -curves for two transformations involving crystals, namely the allotropic transformation of tin and the combination of calcium oxide and ice to form calcium hydroxide. In both cases it will be seen that the curves agree with the requirements of the heat theorem, while the curves for the second case also illustrate the general thermodynamic relations mentioned in section (c). A large number of other reactions between crystals have been investigated thermodynamically and shown to agree with the heat theorem.

Attempts have been made to extend the range of the heat theorem so as to include pure liquids. Although this appears to be justified in some cases, it cannot be regarded as generally applicable. Experimental investigation is rendered difficult by the instability of supercooled liquids at low temperatures, though "amorphous" substances (which in principle may be regarded as liquids) can be investigated in this temperature range.

The theorem does not apply to processes involving gases, as may be seen from the behaviour of their specific heats [III. 2. d.]. Although these specific heats fall with decreasing temperature, if

the translational energy of the molecules behaves classically and is not quantised, all gaseous specific heats must reach a constant minimum value at the lowest temperatures. It may be predicted

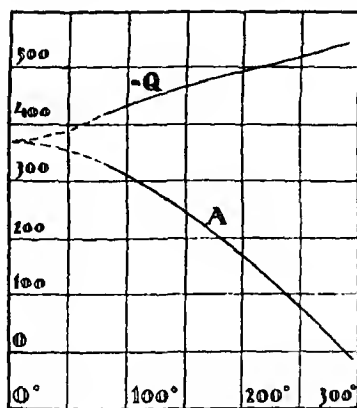


Fig. 3.

that gases will exhibit this classical behaviour at very low pressures. There are, however, reasons for believing that at finite pressures even the translational energy is quantised, thus altering the properties of gases at extremely low temperatures. This pheno-

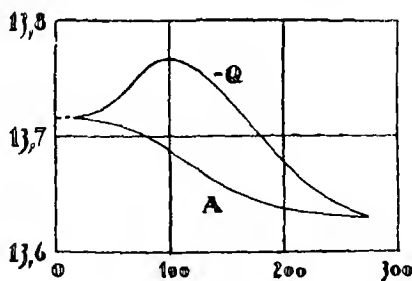


Fig. 4.

menon is termed *gas degeneration*. It has not yet, however, been demonstrated experimentally, and cannot be directly related to the third law.

Finally, as regards the application of the theorem to mixtures, it must be remembered that according to equation IV. (4) the entropy change associated with the formation of a mixture of ideal gases is

independent of the temperature, and hence retains the finite value it possesses at ordinary temperatures right down to absolute zero. It may be assumed that other states of aggregation exhibit similar behaviour, and in particular that the entropy of mixing of crystalline substances retains positive values even at the lowest temperatures. Direct application of the third law to mixtures is hence excluded.

(e) **Thermodynamic Functions and the Third Law.** The considerations of the preceding paragraph refer to affinity and heat of reaction, *i.e.*, to *changes* in the thermodynamic functions brought about by chemical processes. It will simplify calculation to deal instead with the thermodynamic functions of the separate substances involved. For the entropy the appropriate equation has already been given in (11), *i.e.*,

$$S - S_0 = \int_0^T c_p d \ln T.$$

The heat content is given by (20) as

$$H - H_0 = \int_0^T c_p dT, \quad (21)$$

while the equation for the thermodynamic potential is obtained from (8), (11) and (21), putting $S_0 = 0$, giving

$$G - G_0 = \int_0^T c_p dT - T \int_0^T c_p d \ln T. \quad (22)$$

The values of these functions relative to their values at $T = 0$ can hence be calculated at any temperature provided that the values of c_p are known throughout the whole temperature range. In the case of the entropy, Planck's formulation of the third law puts $S_0 = 0$ for pure crystalline substances. From a purely thermodynamic standpoint this postulate goes no further than equation (18), since the absolute values of the thermodynamic functions are without thermodynamic significance. It is however significant from the point of view of statistical mechanics, which is of importance in understanding the concept of entropy. In any case it is a convenient convention for fixing numerical values.

In agreement with this convention we shall write equation (11) for pure crystalline substances in the form

$$S = \int_0^T c_p d \ln T. \quad (23)$$

In order to determine S (or, more generally, $S - S_0$) we may plot c_p against $\ln T$, or c_p/T against T , when the area under the curve obtained will represent the value of the integral. In carrying out this determination it must be remembered that specific heat measurements do not generally extend to the lowest temperatures, but may be replaced or supplemented by the theoretical T^3 -law previously mentioned. Further, the contribution of the range of very low temperatures to the entropy is relatively small provided that no changes of state take place in this range.

If there is a change of state or a polymorphic transformation in the temperature interval considered, the heat content and the entropy will of course be increased by amounts q and q/T respectively, where q is the latent heat and T the transition temperature. At this temperature there will also be a discontinuous change in the direction of the G -curve.

TABLE I. *Thermodynamic Functions for Pb and I*

T	Lead		Iodine	
	S	$G - G_0$	S	$G - G_0$
10	0.23	0.59	0.17	0.4
20	1.29	7.55	0.91	5.3
50	5.05	106	3.86	77.4
100	8.94	465	7.29	365
200	13.08	1588	11.42	1327
300	15.62	3030	13.98	2600

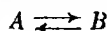
TABLE II. *Thermodynamic Functions for $\frac{1}{2}HCl$*

T	S	$G - G_0$
10	0.045	0.116
20	0.301	1.67
90	3.46	134
100	5.33	173
150	7.44	494
160	9.34	572
180	10.2	767

As an example the above tables give the values of S and $G - G_0$ for lead, iodine and hydrogen chloride. The values at $T = 10^\circ$ are

however somewhat uncertain. For *HCl* the large increases in *S* in the temperature intervals 90°–100° and 150°–160° are due to a polymorphic change and to fusion respectively.

(f) **The Calculation of Gas Equilibria from the Third Law.** For a homogeneous gas equilibrium of the form



we have from VI. (33)

$$\frac{d \ln K_p}{dT} = \frac{H_B - H_A}{RT^2},$$

where $K_p = \frac{p_B}{p_A}$. This gives on integration

$$\ln K_p = \ln \frac{p_B}{p_A} = \int \frac{H_B - H_A}{RT^2} dT + J. \quad (24)$$

In equation (24) and subsequent equations containing an indefinite integral and an integration constant, the integration is to be carried out without introducing any further constants. As previously shown [II. 1. b.], the *H*-functions involved are independent of the pressure but vary with the temperature according to I. (60). *J* is an integration constant which cannot be determined by classical thermodynamics. We shall, however, see that this constant can be determined by recourse to the third law.

For this purpose we introduce P_A and P_B , the vapour pressures of *A* and *B* in the crystalline state. These are given by the integrated form of II. (49), which by analogy with (24) can be written

$$\ln P_B = \int \frac{H'_B - H_B}{RT^2} dT + j_B, \quad (25)$$

$$\ln P_A = \int \frac{H'_A - H_A}{RT^2} dT + j_A, \quad (26)$$

where (as in (24)) H'_B and H'_A refer to the gaseous state. Like *J*, the constants j_A and j_B cannot be determined thermodynamically, but there exists an important relation between these three constants which makes it possible to calculate gas equilibria if the vapour pressure curves are known over a large temperature range.

Combining the three equations (24), (25) and (26) gives

$$\ln \frac{P_A P_B}{P_B P_A} = \int \frac{H_B - H_A}{RT^2} dT + J - (j_B - j_A), \quad (27)$$

or after multiplying by *RT*,

$$RT \ln \frac{P_A P_B}{P_B P_A} = G_A - G_B = T \int \frac{H_B - H_A}{T^2} dT + RT [J - (j_B - j_A)] \quad (28)$$

since the first term is easily seen to be equal to the difference between the G -functions of A and B in the crystalline state. The value of the integral is given by 1. (39) as

$$\int \frac{H_B - H_A}{T^2} dT = -\frac{G_B - G_A}{T} + i_B - i_A. \quad (29)$$

At very low temperatures, however, the third law gives

$$H_B - H_A = G_B - G_A = \text{constant},$$

and if (29) is integrated on this assumption we find

$$i_B - i_A = 0. \quad (30)$$

We can therefore introduce $-\frac{G_B - G_A}{T}$ in place of $\int \frac{H_B - H_A}{T^2} dT$ in equation (28), giving directly

$$J = j_B - j_A.$$

The argument given here for a simple gas equilibrium involving only two components is easily seen to be valid for gas equilibria of any type, so that we can write in general

$$J = \sum j_B - \sum j_A.$$

The introduction of the third law by means of the condition $i_B - i_A = 0$ thus leads to a general method for determining the integration constant J . This constant is found to depend on constants referring to the *separate* substances taking part and on the numbers of molecules involved, but does not depend in any other way on the particular combination considered. The j -constants can be determined from a knowledge of the complete vapour pressure curves of the separate substances involved, and if these constants are known it is possible to make a complete calculation of the homogeneous equilibrium on the basis of the thermal data for the reaction. It will also be seen that calculations can be carried out for homo-heterogeneous reactions in an exactly similar manner, except that the j -constants corresponding to the solid substances are omitted from equation (27).

The vapour pressure constants j are thus of considerable importance for the general solution of the affinity problem, and on account of their use for calculating chemical equilibria they have been termed "*chemical constants*." It is important to note that these constants can in many cases be calculated by means of quantum theory and statistical mechanics from a knowledge of the mass and structure of the molecules in question.

Gas equilibria can be calculated by a formally simpler and more rational method by the direct application of equation I. (30),

$$\Delta G = \Delta H - T\Delta S,$$

to gas reactions in which all the participants are present as pure gases at unit pressure and the same temperature T . By considerations analogous to those used in deriving equation VI. (4), it is easily seen that

$$\Delta G = -RT \ln K_p, \quad (31)$$

so that we can write

$$RT \ln K_p = T\Delta S - \Delta H. \quad (32)$$

In this equation ΔH is the heat of reaction while $\Delta S = S_B - S_A$ is the difference between the entropies of the final and initial systems at unit pressure. These entropies can be calculated on the basis of the third law as described under (c), so that the equilibrium constant K_p can also be determined.

2. ATOMIC AND MOLECULAR PROCESSES

We have so far considered chemical reactions from the standpoint of macroscopic measurements, *e.g.*, thermodynamic measurements of affinity. From the point of view of molecular theory these reactions must be considered as the sum of a large number of single processes involving the individual atoms and molecules of the substances involved. We have already seen that this point of view can be very fruitful in certain fields. It is, however, possible to study atomic and molecular processes by means of methods differing essentially from the thermodynamic methods so far employed. Such methods may lead to an insight into the more detailed mechanism of physico-chemical processes, and may also be used to test the relations which can be shown to exist between the energy involved in single molecular processes and the usual thermodynamic quantities.

In this connection we shall deal not only with molecular transformations which can be expressed by an ordinary chemical equation, but more generally with any discontinuous changes in atoms and molecules. Phenomena of this kind are governed by the laws of the *quantum theory*, and can only be understood in terms of this theory.

The experimental methods available for investigating the energy relations of single atoms and molecules are chiefly optical or electrical, as already mentioned in connection with the photochemical activation of molecules [VIII. 3]. When such experiments are

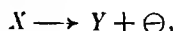
interpreted with the help of quantum theory it is possible to obtain information about the energy states of atoms and molecules, and in particular the amount of work required to pass from one state to another. However, this work cannot be directly identified with the affinity of the macro-reaction, even when the nature of the process admits of such a comparison. This may be seen, for example, by considering the dissociation affinity of a gas. Even if the temperature is constant this affinity will depend on the pressure and in virtue of this dependence may assume both positive and negative values for the same reaction. There can be no such effect of pressure on the quantity referring to a single molecule, since the concept of pressure has no meaning in this connection. This obvious difference is due to the part played by the thermal motion of the molecules in setting up equilibria: the non-uniform distribution of energy (expressed by Maxwell's law) causes certain molecules to have a high kinetic energy which appears as work when dealing with single molecules and leads to the processes which are observed macroscopically. It is, however, clear that this effect of molecular motion will decrease to an unlimited extent with decreasing temperature, so that at absolute zero both the affinity and the heat evolved in the macroscopic process will be identical with the work characteristic of the molecular process.

The meaning of Q_0 for a molecular process also appears from the quantum relation between the energy of this process and a frequency which does not depend on the temperature. For molecular process entirely conditioned by such a relation the distinction between work and energy disappears, and the values for the work obtainable by these methods can be directly compared with the heat evolved by the macro-reaction at absolute zero. Moreover, reactions which can be studied by these methods will in general be characterised by energy changes, which are very large compared with the thermal energy of the order of magnitude kT . It is therefore often unnecessary to refer the heat change to absolute zero, and the more accessible value at the temperature of the experiment may be used.

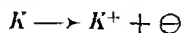
As stated above, these phenomena can only be interpreted on the basis of the quantum theory, according to which not only the energy of atoms and molecules but also the energy of electromagnetic radiation must be regarded as quantised. The old quantum theory involved various inconsistencies: for example, the occurrence of corpuscular properties in light, which in other respects has the character of waves, and the association of wave-like properties with the motion of charged particles. These inconsistencies have to a great extent been fundamentally removed by the new *quantum mechanics* which has developed during the last ten years. We can,

however, do no more than mention the consequent revolution in our fundamental conceptions of simple natural phenomena.

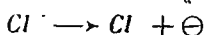
(a) **Work of Ionisation and Electron Affinity.** The work of ionisation is the energy which must be supplied to an isolated particle (charged or uncharged) in order to make it lose an electron. This process may be represented by



so that the work of ionisation for the particle X is the work which must be supplied in order that the process represented by this equation shall take place. The same work is termed the *electron affinity* of the particle Y . This terminology applies independent of the charge on the particles. Thus the energies required for the dissociations



and



represent on the one hand the work of ionisation for the potassium *atom* and the chlorine *ion* respectively, and on the other hand the electron affinity of the potassium *ion* and the chlorine *atom* respectively. It is clear that a high ionisation potential means that the particle is stable in the sense that it has little tendency to change its state, while a large electron affinity corresponds to a large tendency to change.

One method for ionising electrically neutral atoms in the gaseous state is to bombard them with electrons of suitable velocity, the occurrence of ionisation being detected by the power of the gas to conduct electricity (Franck and Hertz).

The principle of the experiment is illustrated in Fig. 5. A tungsten filament T , surrounded by a cylindrical metal electrode R , is heated by

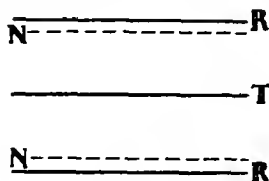


Fig. 5.

an electric current and thus caused to emit slow electrons. R is charged negatively relative to the filament, while the wire gauze N between R and T is charged positively. The electrons are accelerated in the

field TN , and if their initial velocity is zero, their velocity after passing through the field is given by

$$\frac{1}{2}mv^2 = e\Delta\psi. \quad (33)$$

If the potential of the electrode R is sufficiently high the electrons will never reach it, and the increased kinetic energy of the electrons caused by an increase of the potential TN will not appear as an electric current in a galvanometer inserted between T and R . A change in the reading of this instrument only occurs when ionisation is set up, when the positive ions formed are attracted by the electrode R and transfer their charge to it.

Experiment shows that the galvanometer deflection appears suddenly when the accelerating field $\Delta\psi$ reaches a certain critical value. It must be assumed that the electron energy corresponding to this potential is just sufficient to ionise the atoms, so that this method gives the ionisation potential directly.

The work of ionisation is obtained from the ionisation potential by multiplying by the electronic charge, which has the value [IX. 1. b.] 4.774×10^{-10} electrostatic units or 1.591×10^{-19} coulombs. If the charge on the electron is used as the unit of electric charge, so that the electrical energy is expressed in "electron volts," then the numerical value of the ionisation potential is the same as that of the work of ionisation. In this case the "work" can be expressed in volts.

The value of the ionisation potential is usually considerably higher than the potentials met with in the electrochemistry of solutions. The potentials are particularly high for the inert gases, e.g., 24 volts for helium and 16 volts for argon. The ionisation potential of the mercury atom is 10.4 volts, and the values for the alkali metals considerably lower, e.g., 5.1 volts for sodium. This is in agreement with the chemical fact that the alkali metals readily form positive ions, while an analogous phenomenon is unknown for the inert gases under ordinary conditions. It is roughly correct to say that for each group of the periodic system the ionisation potentials increase progressively from the inert gas to the alkali metal.

(b) Resonance Potentials. If in the experiment described in the last section the electrode R is kept at a considerably lower potential, then electrons which are sufficiently accelerated will be able to reach R , producing a deflection in the galvanometer in the opposite direction to that observed before. If this deflection is plotted as a function of the accelerating potential, we obtain not a steadily rising curve but one with several maxima and minima. This is illustrated in Fig. 6, which shows the results for mercury vapour. This phenomenon means that when the accelerating potential exceeds a certain value, the electrons are no longer able to reach the electrode R , though they were able to do so at a slightly lower potential. This can be explained by supposing

that when the velocity of the electrons exceeds a certain critical value, corresponding to the so-called *resonance potential*, they lose energy by colliding with the atoms of the gas, this energy being taken up by the colliding atom. When the velocity of the electron increases above the

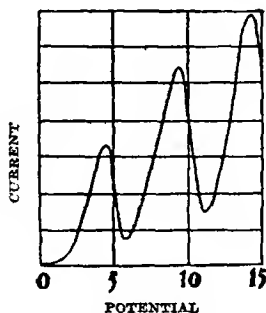


FIG. 6.

critical value the current again rises : this is because the atom only absorbs an amount of energy corresponding to the resonance potential, so that if the energy of the electron exceeds this amount it will continue its motion with a velocity corresponding to the difference between the energies before and after impact. If the potential is increased still further, a second critical point is reached of the same nature as the first, and so on.

This experiment shows that atoms can absorb energy discontinuously in quanta and also gives directly the magnitude of the energy quantum absorbed. The states which the atom assumes after absorbing quanta are known as *excited states*.

(c) **Thermal Ionisation.** The emission of electrons by a hot metal wire (mentioned above) is an example of thermal ionisation, a phenomenon which is exhibited both by solids and by gases at very high temperatures. An electron vapour pressure can be attributed to the metals, and it is a simple matter to calculate the maximum velocity with which electrons evaporate from an incandescent filament by measuring the maximum conductivity between the filament and a surrounding metal electrode.

The atoms of gases dissociate at very high temperatures into electrons and ions, and these equilibria can be treated by thermodynamic methods resembling those used for ordinary chemical equilibria, though it must be noted that the ionic charge has an important effect on the magnitude of the equilibrium constant. The treatment is based on the thermodynamic equations developed in the preceding sections, in particular equation (24). The ionisation energy at absolute zero can be determined from the ionisation

potential as described under (a.), or (more generally) from *optical spectra* (e.). The specific heats needed to integrate (24) are readily obtained for a dissociating metal vapour, since in this case all the participants are monatomic gases for which $c_p = 5$ independent of the temperature. Finally, the value of J in (24) can be calculated by means of quantum theory.

Since the thermal ionisation of metal vapours takes place to an appreciable extent only at very high temperatures, calculations of this kind are particularly important in astrophysics, *e.g.*, for stellar atmospheres. As an example it may be mentioned that the ionisation of the calcium atom to Ca^+ and an electron is calculated to take place to the extent of about 50% at $10,000^\circ$ and a pressure of one atmosphere.

(d) **The Quantum Theory Model of the Atom.** The resonance potentials mentioned above can be accounted for by means of the quantum theory model of the atom, which has been developed largely on the basis of observations on atomic spectra (N. Bohr). It is assumed that atoms contain a nucleus, possessing almost the whole mass of the atom, and having a positive charge, the magnitude of which is given by the number of elementary charges equal to the number representing the position of the atom in the periodic system. In the electrically neutral atom the nucleus is surrounded by an equal number of electrons. The stability of a system of this kind cannot be accounted for on the basis of classical mechanics. In the quantum theory it is assumed that the atom can exist in certain stationary states, between which only discontinuous energy changes can take place. The resonance potentials are the potentials corresponding to the changes in atomic energy associated with these transitions, a quantum of energy being absorbed or emitted. The lowest energy level corresponds to the normal state or *ground state* of the atom, while the excited states are the result of the absorption of quanta of energy by the atom.

It has previously been mentioned [III. 3. b.] that the energy quantum for a linear harmonic oscillator is

$$E = h\nu, \quad (34)$$

where ν is the frequency and h Planck's constant. The energy of the oscillator is hence

$$E = nh\nu, \quad (35)$$

where n can have the values $0, 1, 2, \dots$

* A similar treatment can be applied to a rotator consisting of a particle rotating in a circle. The energy of the oscillator is equally divided between kinetic and potential energy, while the rotator

does not possess any potential energy in virtue of its rotation. Its energy is thus only half as great as that of the oscillator, *i.e.*,

$$E_k = \frac{n}{2} h \nu = \frac{n}{4} h \frac{\omega}{\pi}, \quad (36)$$

where ν is the frequency of rotation and ω the angular velocity. Further, classical mechanics gives for the rotator

$$E_k = \frac{1}{2} J \omega^2, \quad (37)$$

where J is the moment of inertia. This gives

$$\omega = 2\nu\pi = \frac{n\hbar}{2\pi J} \quad (38)$$

and

$$E_k = \frac{n^2 \hbar^2}{8\pi^2 J}. \quad (39)$$

This equation can be applied to the rotation of a diatomic molecule, and hence used to calculate the contribution to the specific heat of diatomic gases arising from the rotational energy.

If the rotator is an atom, *i.e.*, an atomic nucleus with electrons rotating round it, then analogous equations can be derived. The simplest example is the hydrogen atom with a single nuclear charge and one electron in a circular orbit. According to Coulomb's law the force between the two charges is e^2/r^2 , and since this is equal to the centrifugal force, we have

$$\frac{e^2}{r^2} = \frac{m\omega^2}{r}, \quad (40)$$

where m is the mass of the electron and ω its velocity. The potential energy of the atom (*i.e.*, the work necessary to remove the electron) is thus found to be

$$E_p = \int_r^\infty \frac{e^2}{r^2} dr = -\frac{e^2}{r}. \quad (41)$$

We can write for the kinetic energy of the electron

$$E_k = \frac{1}{2} m \omega^2 = \frac{1}{2} m r^2 \omega^2 \quad (42)$$

or by introducing (40),

$$E_k = \frac{e^2}{2r}. \quad (43)$$

Combining this equation with (36) we thus obtain

$$E_k = \frac{1}{2} m r^2 \omega^2 = \frac{e^2}{2r} = \frac{n}{4} h \frac{\omega}{\pi}, \quad (44)$$

and hence for the total energy

$$E = E_p + E_k = -\frac{e^2}{r} + \frac{e^2}{2r} = -\frac{e^2}{2r}. \quad (45)$$

Finally, eliminating $r\omega$ from (44), we find

$$E = -E_k = -2m\left(\frac{e^2\pi}{nh}\right)^2, \quad (46)$$

which gives the energy of the hydrogen atom for different values of n , i.e., for the different stationary quantum states.

(e) **The Relation between Radiation and Atomic Energies.** An atom can be excited not only by collision with moving electrons and by conversion of thermal energy, but also by absorption of radiation. Conversely, the transition from a higher to a lower quantum state can be accompanied either by the transfer of energy to colliding particles, or by the *emission* of radiation. In classical optics it was assumed that the frequency of the radiation coincided with the frequency of the material oscillator or rotator producing the radiation. In quantum theory, on the other hand, it is assumed that the radiation frequency is determined entirely by the energy difference between the two stationary states in which the atom exists. This is expressed by the general *frequency law*,

$$h\nu = E_2 - E_1, \quad (47)$$

where h is Planck's constant, ν the frequency of the radiation and E_2 and E_1 the energies of the two states of the oscillating or rotating particle. The frequency of the motion of the particle thus does not enter directly into the equation, and it will only coincide with ν in particularly simple cases, e.g., the harmonic oscillator.

The frequency law has been used in the theory of *photochemical activation* [VIII. 3. b.] and as a basis for explaining the *photoelectric effect*, which consists in the emission of electrons by metals on illumination with visible or ultra-violet light. It is also concerned in the production of *X-rays* of wavelength 0.1—10 Å when solids are bombarded with very fast electrons. Particular importance attaches to the use of the frequency law in interpreting *optical spectra* (Bohr).

The classical theory of electromagnetic radiation is not consistent with the model of the atom described above, since it predicts a continuous emission of radiation by the atom as long as the electrons circulate round the nucleus, and also that the frequency of the light emitted should be the same as that of the material particle. We shall however see that there is to a certain extent agreement between the latter prediction and the frequency law (47).

If we consider two states with values of n differing by unity, equation (39) gives

$$E_2 - E_1 = \frac{(2n+1)h^2}{8\pi^2 J}, \quad (48)$$

J being assumed constant. Combining this with equation (38) in the form

$$h\nu_r = \frac{nh^2}{4\pi^2 J}, \quad (49)$$

where ν_r is the rotation frequency, we find

$$E_2 - E_1 = \frac{2n+1}{2n} h\nu_r. \quad (50)$$

For large values of n this expression becomes

$$E_2 - E_1 = h\nu_r, \quad (51)$$

in agreement with (47). Thus for large values of n the radiation frequency predicted by classical theory for a rotator (i.e., the rotation frequency ν_r) becomes identical with the frequency calculated from (47). This is one way of expressing the so-called *correspondence principle*.

The energy difference between two stationary states corresponding to $n = n_1$ and $n = n_2$ is given by (46) as

$$E_2 - E_1 = -\frac{2me^4 n^2}{h^3} \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right). \quad (52)$$

Combining this with the frequency law (47) we obtain

$$\nu = -\frac{2me^4 n^2}{h^3} \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right). \quad (53)$$

Further, we have

$$\nu\lambda = c = 3.00 \cdot 10^{10}, \quad (54)$$

where λ is the wavelength corresponding to the frequency ν and c is the velocity of light, giving

$$\frac{1}{\lambda} = -\frac{2n^2 me^4}{ch^3} \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right), \quad (55)$$

or

$$\frac{1}{\lambda} = -R^* \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right), \quad (56)$$

where R^* is known as the Rydberg constant. In order to calculate the value of this constant we take the values

$$\begin{aligned}
 \pi &= 3.142 \\
 m &= 0.903 \cdot 10^{-27} \\
 e &= 4.77 \cdot 10^{-10} \\
 h &= 6.547 \cdot 10^{-27} \\
 c &= 3.00 \cdot 10^{10}
 \end{aligned}$$

whence

$$R^* = \frac{2\pi^2 me^4}{ch^3} = 1.096 \cdot 10^5. \quad (57)$$

The wavelength is hence given by (56) as

$$\frac{1}{\lambda} = -1.096 \cdot 10^5 \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right), \quad (58)$$

where λ is expressed in centimetres. If $n_1 = 2$ and λ is expressed in Ångström units, (58) becomes

$$\lambda = 3650 \frac{n_2^2}{n_2^2 - 4}. \quad (59)$$

The theory based on the frequency condition (47) thus predicts that the hydrogen spectrum should contain lines with wavelengths given by putting $n_2 = 3, 4, 5$, etc., in equation (56). These lines correspond to transitions to the orbit with $n = 2$ from orbits with higher quantum numbers. Such lines are in fact found in the hydrogen spectrum in the so-called Balmer series. Other series which occur in the hydrogen spectrum can be calculated from (55) by putting $n_1 = 1, 3$ or 4 , representing electronic transitions to the corresponding energy levels from higher quantum states.

The energy of the atom in the different quantum states can be calculated from (52) and (57), giving

$$E_2 - E_1 = -R^*hc \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right), \quad (60)$$

or, inserting the numerical values,

$$E_2 - E_1 = -2.16 \cdot 10^{-11} \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right), \quad (61)$$

where the energy is expressed in ergs. The energy in electron-volts is obtained by dividing by 1.591×10^{-12} , which is the ratio between an electro-volt and an erg. This gives

$$E_2 - E_1 = -13.56 \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right). \quad (62)$$

The energy necessary to remove an electron completely from a hydrogen atom in the ground state is obtained from (61) by putting

$n_2 = \infty$, $n_1 = 1$, while the ionisation potential is obtained in the same way from equation (62). This gives respectively 2.16×10^{-11} ergs and 13.56 volts, or, expressing it in calories per gram-molecule, 313,000 cals. The wavelength corresponding to $n_2 = \infty$ can be read off from the spectrum at the point to which the lines in the series converge, and may also be recognised as the limit of the continuous spectrum, corresponding to continuous variations in the kinetic energy of the free electron. In spectra with a more complex structure the ionisation potential can often be determined from the convergence of a line series (*e.g.*, for the alkali metals), and there is satisfactory agreement with the potentials determined by other methods (*a.*).

The numerical relation between radiation frequency and energy has already been given in equation VIII. (46). The relation between frequency and potential is obtained by combining (47) and (54) and using the above value for the energy of an electron-volt, giving

$$\psi\lambda = 1.235 \times 10^{-4}. \quad (63)$$

As an example of the application of this equation we may mention the fact that when mercury vapour is excited by electron bombardment it emits light of wavelength 2536.7Å at the critical potential of 4.9 volts mentioned above (*b.*). Introducing the value

$$\lambda = 2.5367 \times 10^{-5} \text{ cms.}$$

in equation (63) we find

$$\psi = 4.87 \text{ volts}$$

in good agreement with the value found directly.

Conversely, if mercury vapour is illuminated with light of this wavelength it will be excited, and the excited atoms may subsequently produce chemical effects by collision with other particles. For example, they are able to dissociate the hydrogen molecule into atoms (*photo-sensitisation*).

It is thus possible to obtain information about the energy relations of atoms from spectrum observations, and this is particularly important because the wavelengths of spectrum lines can be measured with an accuracy far exceeding that of most other physical measurements.

The visible spectrum extends roughly from 4,000—8,000Å, which according to equation (63) corresponds to potentials between 3 and 1.5 volts. Ultra-violet radiation goes down to wavelengths of about 100Å, and infra-red radiation up to wavelengths of about 10^6 Å, corresponding respectively to higher and lower potentials.

X. 2. ATOMIC AND MOLECULAR PROCESSES 345

The spectra of higher atoms have a more complicated structure than that of the hydrogen atom, but the same fundamental ideas can be applied. The most simple spectra are the so-called spark spectra of light atoms like helium and lithium, in which the emission is due to the quantised transition of a single electron relative to a multiply charged nucleus.

(f) **The Relation between Radiation and Molecular Energies.** A molecule has of course the same sources of energy as an atom, but as already mentioned in [III. 2. d.] it is also able to accumulate energy by the rotation and vibration of the atoms in the molecule. The quantity of these forms of energy can only change by quanta, and these changes can give rise to the emission or absorption of radiation. We should therefore anticipate that molecular spectra will exhibit a higher degree of complexity than atomic spectra.

The rotational energy is given by equation (39),

$$E_r = \frac{n^2 h^2}{8\pi^2 J}.$$

If the rotating molecule possesses a permanent electric moment it will absorb radiation of frequency given by the general frequency equation (47). By combining these two equations we find

$$h\nu = E_{r_2} - E_{r_1} = \frac{h^2}{8\pi^2 J} (n_2^2 - n_1^2), \quad (64)$$

involving the approximately correct assumption that the moment of inertia J is independent of the frequency of rotation. Since in rotational transitions n only changes by one unit, $n_2 = n_1 + 1$, giving

$$\nu = \frac{h}{8\pi^2 J} (2n + 1), \quad (65)$$

with $n = 0, 1, 2$, etc. We thus have for an increase of unity in the quantum number

$$\Delta\nu = \frac{h}{4\pi^2 J}. \quad (66)$$

The rotation spectra of diatomic molecules lie in general in the long-wave infra-red. Measurements of these spectra confirm the even spacing of the lines demanded by (66), *e.g.*, for HCl we have for n -values from 4 to 10

$$\Delta\nu = 6.3 \times 10^{11}.$$

* Equation (66) then gives for the moment of inertia

$$J = 2.64 \times 10^{-40}.$$

Inserting this value in the formula for the moment of inertia,

$$J = \frac{m_1 + m_2}{m_1 m_2} d^2, \quad (67)$$

where m_1 and m_2 are the masses of the two atoms and d the distance between their centres of gravity in the molecule, we find

$$d = 1.28 \times 10^{-8}.$$

The value for the dimensions of the hydrogen chloride molecule derived from the rotation spectrum thus agrees with the values for molecular dimensions obtained by other methods.

The vibrational energy of a molecule is in general characterised by a higher frequency than the rotational energy, and the corresponding spectrum therefore occurs in the short-wave infra-red. If a simultaneous change in both rotational and vibrational states takes place in the molecule, the frequency of the light emitted will be related to the sum of the energy changes by an equation analogous to (47),

$$h\nu = (E''_v - E'_v) + (E''_r - E'_r). \quad (68)$$

Since the vibrational energy E_v is considerably greater than the rotational energy E_r , the spectrum will consist of a series of bands, each band corresponding to a definite difference $E''_v - E'_v$, while the changes in rotational energy appear in each band as a "fine structure." The rotational spectrum is thus in a sense projected into a shorter wave region of the infra-red, which from an experimental point of view makes it more accessible than the pure rotational spectrum.

Analogous considerations apply to the spectrum which is produced by simultaneous changes in the rotational, vibrational and electronic energies, and we can write

$$h\nu = (E''_v - E'_v) + (E''_r - E'_r) + (E''_e - E'_e), \quad (69)$$

where the last term is the energy corresponding to an electronic transition. The spectrum will then consist of a number of band-systems. The position of each system in the spectrum is determined by E_e , the position of the single bands in the system by E_v , and finally the line structure of each band by E_r . The electronic energy E_e is of the same order of magnitude as the electronic energies in atomic spectra.

(g) Energy and Heat of Dissociation. If a diatomic molecule acquires increasing quantities of vibrational energy, it will reach a state in which the amplitude of vibration has become so great that dissociation occurs. The energy necessary for this can be calculated from molecular spectra by means of the quantum theory.

A harmonic oscillator obeys equation (35), where ν is the constant frequency. With increasing quantum number n an actual oscillator will in general deviate from simple harmonic motion. This means that ν is not constant, but can be written

$$\nu = \nu_0 + \alpha n, \quad (70)$$

giving for the energy

$$E = n\hbar\nu = n\hbar(\nu_0 + \alpha n), \quad (71)$$

and for the energy increase in passing from n to $n + 1$,

$$\Delta E = \hbar\nu_0 + \alpha\hbar(2n + 1). \quad (72)$$

Since α is in general negative, the energy increase produced by a charge $\Delta n = 1$ will decrease with increasing value of n . In practice it is found that the energy levels converge to a definite limiting value. When this value is reached any further increase of energy brings about separation of the two atoms, so that the convergence energy and the dissociation energy are equal.

According to equation (47), the convergence energy corresponds to a definite convergence frequency, which (as for the ionisation energy of atoms (e.)) can be determined from the spectrum as the limit of a series, or as the point at which the continuous spectrum begins. The dissociation energy in electron-volts can be obtained by inserting the appropriate wavelength in equation (63).

Thus, for example, the convergence of the lines in the absorption spectrum of molecular iodine at $\lambda = 4995\text{\AA}$ gives for the dissociation energy $\psi = 2.48$ volts, corresponding to 57,200 cal. per mole.

Actually, however, this absorption spectrum is not caused by the dissociation of normal iodine molecules, but of iodine molecules in an excited state. The above value must therefore be diminished by the excitation energy, 22,600 cal., giving 34,600 cal. as the dissociation energy of the normal molecules. This agrees with the value previously calculated from equation VI. (35), showing that the spectrum method is applicable to affinity measurements.

The table below gives the heats of dissociation of various diatomic molecules calculated by analogous methods.

TABLE III. *Heats of Dissociation*

H_2	100000
N_2	207500
O_2	117400
Cl_2	56900
Br_2	45200
I_2	34600
HI	66000
NO	157000

In cases where values are available from direct calorimetric measurements or thermodynamic calculations, they are generally in good

agreement with the spectroscopic values. For example, the heat of dissociation of hydrogen has been determined by various methods (including the direct union of atomic hydrogen, which is fairly stable), giving values close to 100,000 cal. The optical methods have, however, considerable advantages in simplicity and accuracy. In principle the energy change in any chemical reaction whatever can be determined spectroscopically, since the energy change in any process is equal to the difference between the dissociation energies of the two systems taking part.

CHAPTER XI

SURFACE AND COLLOID CHEMISTRY

It has already been mentioned [II. 2. a.] that the surface of a solid or liquid phase, or the boundary between two phases, has properties differing from those relating to the interior of the phases. In general, however, physico-chemical equilibria can be treated without taking this fact into account, since the surface usually constitutes so small a proportion of the whole system that the deviations due to its presence are insignificant. If the magnitude of the surface increases relative to that of the whole system, its properties will of course become of increasing importance for the system as a whole. If the surface is relatively very large (as in porous substances), or if we are dealing with very small quantities of substance in a large volume, then surface phenomena can play a predominant part. The section of physical chemistry dealing with these phenomena is termed *surface chemistry*.

A surface of relatively large extent can be obtained by fine subdivision of one of the phases of the system. For example, by shaking together two liquids it is often possible to produce systems in which fine drops of one liquid are suspended in the other, thus producing a very large surface between the suspended drops and the surrounding medium. The treatment of such systems thus constitutes a branch of surface chemistry. On the other hand, the large surface is associated with the presence of *particles* which are capable of free existence and motion and in many respects can be regarded as large molecules. A large number of special properties and phenomena depend upon this fact, and since such systems occur very frequently both naturally and artificially, it is convenient to treat them as a special division of the subject under the name of *colloid chemistry*. It is, however, clear that there can be no sharp division between surface chemistry on the one hand and colloid chemistry on the other.

I. SURFACE CHEMISTRY

We have previously touched on the phenomena of surface chemistry in connection with the thermodynamic treatment of pure liquids. In the present section we shall consider the laws for

the surface composition and surface tension in systems of several components, though out of the many types of possible equilibria we shall only consider a few which are particularly simple or particularly important.

(a) **The Adsorption of Gases.** If a solid is brought into contact with a gas, molecules of the gas will in general be held fast on the surface of the solid. This phenomenon shows that there must be attractive forces between the molecules of the solid and those of the gas, and if the extent of the adsorption is great these forces must be greater than the forces between two molecules of the gas in contact. In general, the forces which bring about adsorption may be regarded as identical with the ordinary van der Waals' cohesive forces, which are also effective between like molecules and are responsible for the simple condensation phenomena. In many cases, however, the attraction is of a more intimate nature and must be ascribed to chemical "valency forces."

Since adsorption is a phenomenon characteristic of surfaces or interfaces, it occurs to a marked degree with porous substances with a large surface, of which a large number are known. Among the adsorbing substances or *adsorbents* in common use we may mention various forms of carbon, such as charcoal and animal charcoal, and silicic acid or aluminium hydroxide in a partially hydrated state. The relative surface area of these substances may reach very high values, e.g., varieties of charcoal are known having a surface of about 1,000 sq. metres per gram. As regards the part played by the adsorbed substance, or the *adsorbate*, it is in general true that gases which are readily condensable are adsorbed to a greater extent than gases which are difficult to liquefy. There are, however, many exceptions to this rule, since when the adsorption depends upon chemical forces the individual chemical affinities will become important. For example, gases like oxygen and carbon monoxide are often strongly adsorbed by metal surfaces. Further, the nature of the surface of the adsorbent is of importance: thus water vapour is adsorbed to a greater extent on a corroded glass surface than on a fresh one.

In its simplest form the phenomenon of adsorption is characterised by the establishment of a reversible equilibrium state, the *adsorption equilibrium*, in which under given conditions the amount of gas adsorbed increases and decreases continuously with the pressure. This equilibrium is generally set up rapidly. In exceptional cases an equilibrium is set up slowly, and this indicates as a rule that the gas molecules are passing into the interior of the adsorbent, in which case the phenomenon is more generally described as *sorption*. The adsorption equilibrium will of course depend upon

the temperature and the pressure (in addition to the chemical and physical nature of the substances concerned), and one of the most important problems in surface chemistry is to establish the laws governing this dependence.

The first requisite is a basis on which to define the amount adsorbed in a given case. In dealing with the adsorption of gases this does not in general involve any difficulty, since either the increase in the weight of the adsorbent or the decrease in the pressure of the gas when the equilibrium is set up can be used directly to determine the adsorption equilibrium. For a given quantity of adsorbent of given composition and structure, the amount adsorbed at equilibrium at a given temperature will depend upon the pressure or concentration of the gas. The equation relating the amount adsorbed y to the equilibrium concentration c is known as the *adsorption isotherm*. In many cases it assumes the form

$$y = kc^{\nu} \quad (1)$$

where k and ν are constants depending on the nature of the substances. This equation is shown graphically in Fig. 1, and is

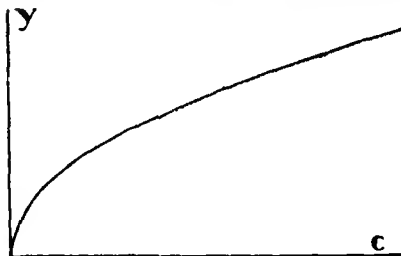


Fig. 1.

commonly known as the Freundlich adsorption isotherm. The exponent ν is a proper fraction (generally greater than 0.2), so that the amount adsorbed increases with increasing pressure, but is not directly proportional to it. In this respect there is a sharp distinction between the phenomena of adsorption and absorption. The Freundlich equation is, however, only valid over a fairly limited pressure range, and in particular it must fail when the pressure is very small.

For a given equilibrium concentration, the amount of substance adsorbed y is of course proportional to the extent of the adsorbing surface. In characterising the adsorption equilibrium it would therefore be rational to calculate y as the amount of adsorbate per

unit area of surface. As a rule, however, this calculation is not possible, since to obtain an adsorption which can be easily measured it is necessary to use porous materials having a surface per unit area which is very large but generally unknown. In such cases y is defined as the amount adsorbed per unit mass of adsorbent, and to obtain comparable figures it is of course necessary that the adsorbing material shall be uniform throughout.

In those cases where measurements have been carried out with a surface of known extent it has been found that the amount adsorbed often corresponds to a single layer of molecules of the adsorbate on the surface of the adsorbent. In the case of a "unimolecular" layer of this kind kinetic considerations make it possible to derive an equation having more theoretical foundation than the empirical Freundlich adsorption isotherm. We assume that each square centimetre of the surface can accommodate a maximum number of gas molecules n_0 , so that we can imagine unit area as containing n_0 places on which a gas molecule can be attached. It is further assumed that the probability (averaged over a period of time) that the collision of a gas molecule with the surface results in adsorption is proportional to the concentration of gas molecules and also to the number of vacant places per unit area. The velocity of condensation in the equilibrium state can then be written in the form

$$k_1 = k_1(n_0 - n)c$$

where n is the number of places per unit area occupied by adsorbed molecules, c the concentration in the gas phase, and k_1 a constant depending on the temperature and the nature of the substances concerned. If it is now assumed that the probability of evaporation (or desorption) of a molecule from the surface is proportional to the number of adsorbed molecules, the desorption velocity can be written as

$$k_2 = k_2 n.$$

At equilibrium k_1 and k_2 must be equal, so that

$$k_1(n_0 - n)c = k_2 n. \quad (2)$$

If we introduce the "degree of saturation"

$$\alpha = \frac{n}{n_0} \quad (3)$$

and also write

$$\frac{k_2}{k_1} = K$$

we obtain the following equation,

$$\alpha = \frac{c}{c + K} \quad (4)$$

In this equation, which is known as Langmuir's isotherm, the concentration can of course be replaced by the pressure without affecting the form of the equation. According to equation (1), the adsorption curve should be horizontal at $c = 0$ and the adsorption should increase without limit as the concentration of the adsorbate increases. According to equation (4), on the other hand, for small values of c the amount adsorbed should be proportional to the concentration, and at high concentrations it should approach asymptotically to a saturation value, corresponding to complete covering of the surface of the adsorbent by a unimolecular layer. Increase of temperature will lead to an increase in k_2 (as for the velocity of vaporisation) and presumably a decrease in k_1 . The amount adsorbed according to equation (4) will hence be smaller at higher temperatures, in agreement with experiment. Over a certain intermediate concentration interval there will not be much difference between the predictions of equations (1) and (4), so that the experimental material used to verify equation (1) can in many cases be used equally well to verify equation (4).

The linear relation between α and c for small concentrations is equivalent to applying Henry's law to adsorption equilibria. Over a larger concentration range it is simpler in graphical treatment to use $\alpha/(1 - \alpha)$ in place of α , since equation (4) shows that there is a linear relation between the former and c over the whole concentration range.

If we now introduce equation IV. (12), it follows that the chemical potential of an adsorbate following Langmuir's equation can be written in the form

$$\mu = RT \ln \frac{\alpha}{1 - \alpha} + \mu_{0.5},$$

where the last term is the potential for $\alpha = 0.5$, i.e., for half-saturation.

Equation (4) is based on the assumption that a unimolecular layer of adsorbed gas molecules is formed on the surface of the solid. This assumption is not unreasonable in view of the well-founded belief that the range of intermolecular attractions is extremely short. Thus it may well be supposed that the solid surface can hold those gas molecules with which it is in direct contact, but not those which are distant one molecular diameter from the surface. On the other hand, it is not out of the question

that the attraction may be effective over somewhat larger distances, e.g., by polarisation of the adsorbed molecules. In this way several layers of molecules may be adsorbed, though less firmly, and the system will deviate from the adsorption isotherm (4) derived above.

In order to calculate the pressure change which can be produced by the adsorption of a unimolecular layer of gas molecules, we shall imagine that the adsorption takes place on the six faces of a cube with sides 1 cm, filled with gas at a pressure of one atmosphere. The total number of molecules present at ordinary temperatures is easily calculated from Avogadro's number to be about 3×10^{19} . If the molecular diameter is 3×10^{-8} cm., a tightly packed unimolecular layer will contain about 10^{15} molecules per sq. cm., so that there are about 6×10^{15} molecules on all six faces. The fraction adsorbed is thus 2×10^{-4} , which is only just measurable. On account of the form of the adsorption isotherm, conditions will be more favourable to experimental measurement at lower pressures, but these figures show clearly that in general accurate measurements on adsorption phenomena can only be carried out by using porous adsorbents with a very much greater surface.

In certain cases, e.g., when water vapour is held by a glass surface, the amount retained is easily measured by weighing. If the thickness of the layer is calculated it is found to be many times greater than that of a unimolecular layer. As already mentioned, it is assumed in such cases that we are dealing with a *sorption* phenomenon, and that the molecules of the vapour pass into the interior of the glass.

As stated above, an increase of temperature will decrease the adsorption. It hence follows that the decrease in adsorption is accompanied by an evolution of heat, since equation II. (49), derived for the ordinary vaporisation process, must also be applicable in the present case. The *heat of adsorption* can be determined either by direct calorimetric measurements, or by applying this equation to the experimental adsorption isotherm at different temperatures. In cases where adsorption involves the formation of a definite chemical link, the heat of adsorption can reach very considerable values. For example, when carbon monoxide is adsorbed on copper there is a heat evolution of 30,000 cal. per gram molecule.

In agreement with this temperature dependence, the phenomenon of adsorption is particularly prominent at low temperatures. Thus by using porous forms of carbon (blood charcoal, coconut charcoal) cooled in liquid air it is possible to produce extremely high vacua.

(b) **Adsorption in Solution.** On the basis of Henry's law for the distribution of a substance between a gaseous and a liquid phase, it seems reasonable to assume that the laws valid for the adsorption of gases will also be valid in a general way when the substance

adsorbed is present in solution. It should, however, be noted that the solvent itself will generally be adsorbed on the surface of the adsorbent, so that conditions are not so simple for adsorption from solution as they are for gaseous adsorption. It must be supposed that the surface of the adsorbent is completely covered by solvent molecules, so that the solute has not free access to the surface as in gaseous adsorption. However, as long as the solution is dilute (and the concentration and activity of the solvent thus practically constant) we shall not expect that this will lead to any change in the form of the adsorption isotherm, since the presence of the solvent can be regarded as a constant factor tending to prevent adsorption of the solute. It has in fact been found that the Freundlich adsorption equation is in general applicable to solutions, at least over a certain concentration range. Langmuir's equation has been verified to a smaller extent for solutions; this is partly because the condition of unimolecular adsorption is less often fulfilled for solutions than for gases, so that in many cases it is necessary to assume that quite thick adsorbed layers can be formed from solutions. Moreover, the other conditions on which Langmuir's equation is based are not generally strictly fulfilled in solutions.

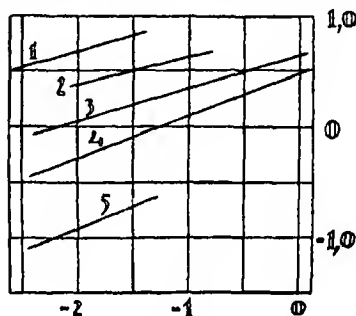


Fig 2

Fig 2 shows graphically some adsorption data in solutions. The graph represents equation (1) in the logarithmic form

$$\log y = v \log c + \log k \quad (5)$$

where y is the number of millimoles adsorbed by 1 gram of adsorbent, and c is the concentration in moles per litre. The substances represented in the five curves are

	Adsorbate	Solvent	Adsorbent
1	Bromine	Water	Blood charcoal.
2	Isoamyl alcohol	Water	Blood charcoal.
3	Succinic acid	Water	Blood charcoal.
4	Benzoic acid	Benzene	Blood charcoal.
5	Picric acid	Ethyl alcohol	Silk

The validity of (5) is shown by the linear relation found to exist between $\log y$ and $\log c$. The exponent ν of the Freundlich equation is found from the slope of the various straight lines. For a large number of solutes and adsorbents of various kinds ν is found to have values between 0.3 and 0.5.

As might be expected, the value of the quantity k , which chiefly determines the adsorption, varies greatly with the chemical nature of the substances taking part in the equilibrium, including the solvent. In adsorption from solutions the solubility of the adsorbate plays a similar rôle to the volatility in gas adsorption. This fact emerges most clearly by considering the effect of the solvent on the adsorption of a given adsorbate by a given adsorbent. It is found as a rule that the amount adsorbed increases with decreasing solubility of the adsorbate in the medium in question, or, in other words, with increasing values of the absolute activity coefficient. If matters were not complicated by simultaneous adsorption of the solvent this rule (and the related rule that ν should be independent of the solvent) would be strictly deducible on a thermodynamic basis.

It is also possible to relate the extent of adsorption to the effect of the adsorbate on the surface tension at the interface between the adsorbent and the solvent, as described under (c)

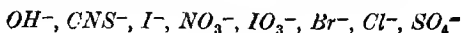
If we compare the adsorption of gaseous and dissolved substances upon a given adsorbent, it is found in general that for the same value of c the amount adsorbed is much greater for gaseous than for dissolved substances. This is partly due to the circumstance mentioned above that the adsorbent in a solution is protected by adsorption of solvent molecules. Further, as in any heterogeneous equilibrium (e.g., simple solubility equilibria) a determining factor is the absolute activity coefficient of the solute, which is normally smaller for dissolved than for gaseous substances. If this is the case the two effects reinforce one another.

As regards the adsorbing properties of different adsorbents, for the reasons mentioned above it is not generally possible to make a rational comparison on the basis of quantity adsorbed per unit area. The order of increasing adsorption for a series of adsorbates is often independent of the adsorbent, so that it is to a certain extent permissible to characterise the single adsorbents and adsorbates independent of the particular way in which they are combined. However, this rule is only approximate owing to the specific effects already mentioned.

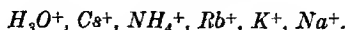
So far we have been dealing chiefly with non-electrolytes or weak electrolytes. Strong electrolytes also as a rule obey the Freundlich adsorption isotherm approximately, but the exponent ν is as a rule considerably smaller than in the case of non-electrolytes, values between 0.1 and 0.2 being not uncommon. The absolute adsorption is also small for typical strong electrolytes unless they contain a

large organic ion, as is the case for many dyestuffs. A substance like crystal violet, $[C(C_6H_4NMe_2)_3]Cl$, is adsorbed to a very marked extent, so that an aqueous solution of it will undergo a readily detectable change in concentration merely on coming into contact with a glass surface.

By comparing salts with a common cation or a common anion, the order of adsorption for different anions or cations respectively can be determined. Using charcoal as adsorbent, the following series have been found, in which the ions are arranged in order of decreasing adsorption tendency,



and



Some special features in the adsorption of salts depend upon the fact that the ions of a salt will in general have different adsorption tendencies, while on the other hand the actual amounts of the ions adsorbed cannot deviate from equivalence to a measurable extent. This causes the adsorption of salts to be accompanied by the formation of an electric double layer, as is in general produced when ions are distributed between two phases. This phenomenon is of great importance in colloid chemistry. Further, the adsorption will be subject to the ordinary mass action effects in the presence of common ions. There may also be changes in the acidity of the solution on the addition of aprotic salts, e.g., if the cation is adsorbed more strongly than the anion the hydroxyl ions from the water will to some extent replace the anion of the salt in the adsorbate and the liquid will acquire an acid reaction. Conversely, an alkaline reaction will be produced if the anion of the salt is more strongly adsorbed than the cation.

If the solution contains several salts an adsorption equilibrium will be set up in which the quantities of the ions adsorbed will depend upon the activities of all the ions present in the solution and upon the specific tendencies to adsorption. This "displacement equilibrium" is a general adsorption phenomenon and is of practical importance in the use of the so-called "adsorption indicators." In this case one of adsorbed substances is coloured, and small changes in the ionic concentrations in the solution can bring about adsorption or desorption of this coloured substance, accompanied by striking colour changes. This type of equilibrium is particularly simple when the total quantity of ions associated with a given quantity of adsorbent remains constant, when a change in the salt-concentrations will only lead to *ionic exchange*. The most important example of this is met with in the so-called *permutites*.

Ordinary permutite is a double silicate of aluminum and either sodium or potassium, having a porous structure. If it is treated with

solutions containing other cations these exchange readily with the alkali metal ions without affecting the structure or outer form of the substance. We may picture the permutite as consisting of a rigid anion skeleton with which are associated mobile cations. Since the charge on the skeleton is constant for a given amount of substance, the number of mobile equivalents of metal is also fixed.

If, for example, sodium permutite is treated with a solution of silver nitrate the sodium ions in the permutite will be partly displaced by silver ions, and an adsorption equilibrium will be set up in which both ions are present in the solution and in the solid phase. The equilibrium conditions can be derived by considerations similar to those used in deriving equation (4). If the concentrations of the two ions in solution are c_1 and c_2 and the quantities of the ions in the permutite n_1 and n_2 , then the velocity of condensation of sodium ions can be written

$$h_1' = k_1 c_1 n_2 = h_2'', \quad (6)$$

since we may assume that it is proportional to c_1 and to the number of places on the adsorbing silicate skeleton which are not occupied by Na . This condensation velocity must be equal to the desorption velocity of the silver ions, h_2'' , since the process is an ionic exchange. Analogously we have

$$h_2' = k_2 c_2 n_1 = h_1'', \quad (7)$$

whence for equilibrium

$$\frac{n_1}{n_2} = K \frac{c_1}{c_2} \quad (8)$$

where K is a constant.

This simple equation has in some cases been confirmed by investigations of ionic exchange in permutite systems, but in most cases there are deviations, showing that the assumptions on which the derivation of (8) is based are not justified. This equation may also be regarded as a consequence of the distribution law V. (31), the permutite representing a phase in which the two ions are present as an ideal mixture, i.e., with activity coefficients which are independent of the composition. However, on account of the small distances between the ions in permutite we can hardly expect these ideal conditions to be satisfied in general, and it is therefore not surprising that deviations from equation (8) are met with in practice.

(c) The Thermodynamics of Interfaces. Gibbs' Adsorption Equation. The term adsorption has so far been used to describe the accumulation of matter which in many cases takes place to a striking extent at the interface between a solid adsorbent and a gas or a solution. Equilibria of this kind have been particularly important in the development of surface chemistry because they are most readily studied experimentally and hence have been much investigated. However, the term adsorption is also used to include a more general phenomenon, namely the fact that the composition of the interior of a phase differs from the composition

of its surfaces or interfaces. The cases so far treated deal with the accumulation of matter, or *positive adsorption*, but in dealing with the more general phenomenon we must anticipate also the occurrence of *negative adsorption*, i.e., a decrease in the concentration of substance at the interface. These various possibilities are related thermodynamically to the surface tension and its variation with concentration, and it is therefore desirable to consider the theoretical treatment of fluid systems in which the surface tension can be measured directly, although an accurate measurement of the actual adsorption in such systems is often difficult.

It has been already pointed out that the boundary between two phases (or between a phase and a vacuum) is not a sharp one, but that there exists a boundary layer in which the composition and properties of the system change continuously in the direction at right angles to the surface. The part of the system containing this layer will be referred to as the *boundary phase*, or *capillary phase*; it should, however, be noted that we have not yet fixed the limits of the boundary phase, and that the meaning of the word phase in this connection differs from that employed in the phase rule [IV. 6], where a phase was assumed to be homogeneous and of indefinite extent. In the boundary phase between two incompletely miscible liquids the continuous concentration change is due to the unsymmetrical environment of the molecules in the layer, and is only present because the thickness of this layer is of the same order of magnitude as the range of the forces of intermolecular attraction. In order to distinguish the boundary phases from the homogeneous phases between which they are formed, we shall term the latter *macro-phases* or *bulk phases*.

In order to arrive at a definition of the composition of a boundary phase it is necessary to define the limits which divide it from the surrounding homogeneous bulk phases. Since the transition between these phases is continuous, there is no *natural* boundary to the capillary phase, and it is not possible to define logically either the thickness of this phase or its concentration (in the ordinary sense of the word). If however we confine ourselves to a qualitative treatment of the question, without demanding any quantitative definitions, then it is possible to compare concentrations without ambiguity. For example, if a capillary phase is said to contain relatively more of a component K_1 than a given bulk phase, this statement will be independent of the position arbitrarily chosen for the boundary between the capillary phase and the bulk phase in question. On this basis it is possible to speak of the "composition" or "concentration" of a capillary phase, and these concepts are sufficiently well defined to lead to a qualitative formulation of

the above-mentioned thermodynamic relation between the surface tension and the concentration.

In order to derive this relation we will begin by considering a binary liquid mixture having a surface tension γ , which will of course vary continuously with the composition of the liquid. We may also assume that the composition of the capillary phase at the surface will in general differ from the composition in the interior of the liquid. We will suppose that the relation between the surface tension and the composition of the bulk of the liquid is represented by the curve $\gamma_2 M \gamma_1$ in Fig. 3, where the ordinates represent surface tensions and the abscissæ compositions, expressed as the mole fraction x .

In order to determine the composition of the boundary phase, we imagine a certain quantity of liquid spread out in a thin film, so

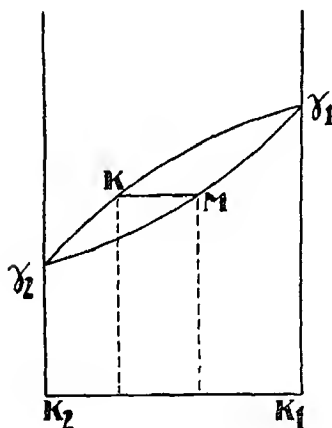


Fig. 3.

that the contracting force due to surface tension is balanced by external forces. In contrast to the case of a pure liquid, a change in the extent of the surface will alter the surface tension, since it alters the composition of the liquid. By an argument closely resembling that previously used for the liquid-vapour equilibrium [IV. 3. g.], it is easily seen that the equilibrium will only be stable if an increase of the surface is accompanied by an increase in the contracting force, i.e., an increase in the surface tension. However, according to the figure an increase in the surface tension is equiva-

lent to an increase in the concentration of component K_1 in the surface, and since this increase is brought about by an increase of surface (i.e., the conversion of macro-phase to capillary phase), the macro-phase must contain relatively more K_1 than the capillary phase. Since the compositions of the two phases change in the same direction, the composition of the mixture transferred must lie between the compositions of the liquid and of the surface, and the thinner the surface layer chosen, the more closely will the composition of the mixture transferred approach that of the surface (cf. the considerations of [IV. 3. g.]). The point K_1 representing the composition of the surface, must therefore lie to the left of point M ,

representing the composition of the liquid, and similarly the whole of the curve $\gamma_2 K \gamma_1$ must lie to the left of the curve $\gamma_2 M \gamma_1$. We thus arrive at the result that the component the addition of which increases the surface tension will be present at a higher concentration in the macro-phase than in the capillary phase. At a point where γ is independent of the concentration (corresponding to a maximum or a minimum in the γ - x -curve) the compositions of the macro-phase and the capillary phase must be identical. These rules are analogous to Konowálow's rule for the liquid-vapour equilibrium, and are a qualitative expression of Gibbs' adsorption equation.

From the point of view of molecular theory, this result may be interpreted as follows. Suppose we add a solute to a given solvent, and that in the mixture the solute molecules are surrounded by a stronger attractive field than the solvent molecules. The presence of some of these solute molecules in the surface will increase the surface tension, since we have seen [III. 2. e.] that the surface tension is determined by intermolecular attractions. The stronger attractive field will also tend to draw the solute molecules out of the surface into the bulk of the liquid. This tendency is opposed by the thermal agitation of the molecules, and we shall therefore have an equilibrium state in which the increase of surface tension is accompanied by a smaller relative concentration of solute in the surface.

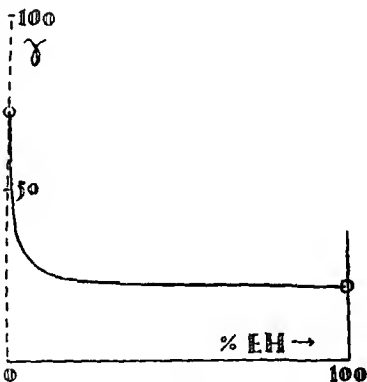


Fig. 4.

Solutes which increase the surface tension of a solvent are often termed *capillary inactive*, while those which decrease the surface tension are termed *capillary active*. This terminology depends upon the fact that an increase of γ is only possible to a limited extent, since molecules tending to produce such an increase will be drawn away from the surface: on the other hand, large decreases in γ are possible, since molecules producing such a decrease will accumulate in the surface, where they are able to produce an effect. As a typical illustration of this effect we may cite the curve, for the surface tension of mixtures of water and acetic acid, shown in Fig. 4.

In order to obtain a quantitative expression for this qualitative

relation it is first necessary to formulate the general thermodynamic equations in such a way as to take into account the surface energy. Thus in equation I. (40) no account is taken of energy changes caused by an increase of surface. Since the increase of surface energy is $\gamma d\sigma$, the complete energy equation for a liquid and its surface should be written

$$dE = TdS - pdv + \gamma d\sigma + \sum \mu_1 dn_1, \quad (9)$$

and the relations for the other thermodynamic functions in I. (44) must be correspondingly modified. Thus for the G-function we have

$$dG = -SdT + vdp + \gamma d\sigma + \sum \mu_1 dn_1. \quad (10)$$

Integration of this equation at constant T , p , γ and μ gives

$$G = \gamma\sigma + \sum \mu_1 n_1, \quad (11)$$

whence by complete differentiation

$$dG = \gamma d\sigma + \sigma d\gamma + \sum \mu_1 dn_1 + \sum n_1 d\mu_1. \quad (12)$$

By combining this equation with (10) we obtain

$$SdT - vdp + \sigma d\gamma + \sum n_1 d\mu_1 = 0, \quad (13)$$

and finally at constant temperature and pressure,

$$\sigma d\gamma + \sum n_1 d\mu_1 = 0. \quad (14)$$

Equation (14) can be applied to any liquid mixture bounded by a surface, and hence to an arbitrary amount of the capillary phase. If the mixture is binary it assumes the form

$$\sigma d\gamma + n_1 d\mu_1 + n_2 d\mu_2 = 0, \quad (15)$$

while for the corresponding bulk phase the ordinary Gibbs-Duhem equation IV. (48),

$$n'_1 d\mu_1 + n'_2 d\mu_2 = 0, \quad (16)$$

must hold. At equilibrium the chemical potentials and their variation with transfer of substance or change of surface must be the same throughout the system. The general equilibrium condition is therefore obtained by multiplying (16) by an arbitrary number x

and adding it to (15), giving

$$\sigma d\gamma + (n_1 - xn'_1) d\mu_1 + (n_2 - xn'_2) d\mu_2 = 0. \quad (17)$$

If now x is chosen so that $n_2 = xn'_2$, (17) reduces to

$$\frac{d\gamma}{d\mu_1} = \frac{xn'_1 - n_1}{\sigma}, \quad (18)$$

which is the Gibbs' adsorption equation in the form applicable to the surface of a single liquid phase.

The quantities involved in equation (18) are completely defined in its derivation, but will be further explained for the sake of clarity. $d\gamma$ is the change in the surface tension of a binary mixture caused by a change $d\mu_1$ in the chemical potential of the first component. This change $d\mu_1$ can be caused either by a change in the absolute amounts of the two components, or by the conversion of macro-phase to capillary phase (or *vice versa*). xn'_1 and n_1 are the quantities of the first component present in the macro-phase and the capillary phase respectively, together with an arbitrary amount n_2 of the second component. σ is the area of the surface. For a given surface and a given composition n_1/n_2 it can easily be shown that the difference $xn'_1 - n_1$ is independent of n_2 , the quantity of the second component taken. In the first place a change in the amount of K_2 at constant surface will transfer equal quantities of K_1 to the macro-phase and to the capillary phase, so that the difference between these quantities must remain constant. In the second place, the left-hand side of equation (18) contains only intensive quantities and must therefore be independent of the arbitrary amount of phase taken; hence the difference $xn'_1 - n_1$ on the right-hand side of the equation must also be independent of the amount of K_2 . This result is of considerable importance in surface chemistry, since it shows that in spite of the inapplicability of the usual concept of concentration, there exists a *completely defined measure of the composition of a capillary phase relative to that of the bulk phase*.

In agreement with the terminology previously used, the difference $n_1 - xn'_1$ is termed the quantity of the substance K , adsorbed by the surface σ . The difference per sq. cm.

$$\frac{n_1 - xn'_1}{\sigma}$$

is termed the surface density, or "surface concentration," of the substance K_1 . Since it is a quantity of substance divided by an area, it is clearly not a concentration in the ordinary sense of the

word. As already mentioned, the adsorption can be either positive or negative. It is easily seen that equation (18) agrees with the qualitative formulation of the law already given.

If the surface density is written as Γ , (18) assumes the simple form

$$\frac{d\gamma}{d\mu_1} = -\Gamma. \quad (19)$$

If we further assume that the bulk phase is so dilute with respect to K_1 that the gas laws are applicable, this equation can be written in the form

$$\frac{d\gamma}{d\ln c} = -RT\Gamma. \quad (20)$$

As appears from its derivation, (18) applies to the surface of a liquid phase which is in equilibrium with a vacuum, or with a vapour phase which is so dilute that it can play no appreciable part in building up the capillary phase. If on the other hand the interface divides two phases of comparable density (*e.g.*, two liquid phases, or a liquid phase and a vapour phase with a high equilibrium pressure), then (18) is no longer valid. This follows from the fact that it only contains quantities referring to the capillary phase and to one liquid phase, while in the case of a system consisting of two partially miscible layers considerations of symmetry demand that the properties of both layers shall be equally represented in the equation.

If we now consider the equilibrium between two incompletely miscible liquids, it is clear from the previous treatment [IV. 3. i.] that at constant temperature and pressure the nature of the two components will determine not only the compositions of the two layers but also the properties of the interphase. An increase of surface will therefore cause no change whatever in the concentrations or the surface tension, in contrast to the case dealt with above. When the surface is increased, with a consequent conversion of the bulk phases to capillary phase, the two liquid layers will contribute to building up the capillary phase in a ratio which is completely determined once the arbitrary boundaries of the capillary phase have been fixed. If these boundaries are chosen so that the amounts of the two components in the capillary phase are n_1 and n_2 , while the corresponding amounts in the two macro-phases are n_1' , n_2' and n_1'' , n_2'' , then it will always be possible to find two numbers x' and x'' , so that

$$\left. \begin{aligned} x'n_1' + x''n_1'' &= n_1, \\ x'n_2' + x''n_2'' &= n_2, \end{aligned} \right\} \quad (21)$$

showing that the capillary phase can be obtained by mixing x' of the first macro-phase and x'' of the second macro-phase. One of these two numbers may be negative.

Since the properties of a two-phase two-component system are thus invariant, a relation corresponding to equation (18) can only be obtained in such systems if a third component is present. If n_1, n_2 and n_3 are the molecular amounts of the components K_1, K_2 and K_3 in the capillary phase, and $n_1', n_2', n_3'; n_1'', n_2'', n_3''$ the corresponding molecular quantities in the two macro-phases, then application of (14) and IV. (49) gives the following equations, analogous to those already obtained,

$$\sigma d\gamma + n_1 d\mu_1 + n_2 d\mu_2 + n_3 d\mu_3 = 0, \quad (22)$$

$$n_1' d\mu_1 + n_2' d\mu_2 + n_3' d\mu_3 = 0, \quad (23)$$

$$n_1'' d\mu_1 + n_2'' d\mu_2 + n_3'' d\mu_3 = 0. \quad (24)$$

If (23) and (24) are multiplied respectively by x' and x'' , and the resulting equations subtracted from (22), we obtain

$$\left. \begin{aligned} \sigma d\gamma + [n_1 - (x'n_1' + x''n_1'')] d\mu_1 \\ + [n_2 - (x'n_2' + x''n_2'')] d\mu_2 \\ + [n_3 - (x'n_3' + x''n_3'')] d\mu_3 \end{aligned} \right\} = 0, \quad (25)$$

If now the arbitrary numbers x' and x'' are chosen in such a way that (21) holds for the components K_2 and K_3 , then we have

$$\frac{d\gamma}{d\mu_1} = \frac{(x'n_1' + x''n_1'') - n_1}{\sigma}. \quad (26)$$

This equation is exactly analogous to equation (18), and in particular the significance of the quantity in brackets on the right-hand side is exactly analogous to the significance of the term xn_1' in (18). The quantities n_1 and $x'n_1' + x''n_1''$ represent the amounts of K_1 associated with the amounts n_2K_2 and n_3K_3 in the capillary phase and the macro-phases respectively. By analogy with the preceding case $n_1 - (x'n_1' + x''n_1'')$ can be termed the amount of K_1 adsorbed, and the quantity

$$\frac{n_1 - (x'n_1' + x''n_1'')}{\sigma} = \Gamma$$

is termed the surface density or surface concentration. On introducing the last expression (26) assumes the same form as (19). Considerations analogous to those employed in the previous case

show that this surface density is independent of the boundaries chosen for the capillary phase, and hence independent of x' and x'' . Similarly, if the gas laws are obeyed, (26) can be written as (20).

All the quantities in (18) and (26) are measurable in principle, but the quantity adsorbed is usually very small compared with the other quantities involved, and a direct verification of the Gibbs' adsorption equation involves considerable practical difficulties.

(d) **Surface Films.** If a small quantity of liquid is added to a medium in which the liquid is insoluble, when equilibrium is reached the liquid is usually present as a drop, or some other aggregate in which the three linear dimensions are of the same order of magnitude. This is a consequence of the positive surface tension existing between immiscible liquids, resulting in a tendency for the extent of the surface to decrease. In certain cases, however, when such an experiment is carried out it is found that the added liquid spreads out on the surface of the medium as a thin layer of very great area. This phenomenon is observed when oleic acid is added to water. It may be accounted for by the following considerations.

The formation of a large interface depends on an attraction between the medium and the molecules of the added liquid; in the example cited an attraction between water and the carboxyl groups in the oleic acid molecules. An attraction of this kind will normally lead to complete miscibility, but in this case the long paraffin chain attached to the carboxyl group is not appreciably attracted by the medium, but tends (like the molecules of the higher hydrocarbons in general) to form an independent layer on the surface of the water. We must therefore imagine the surface film formed to be *oriented*, in the sense that all the molecules point in the same direction with the carboxyl groups immersed in the water and surrounded by water molecules, while the paraffin chains project above the surface of the water. According to this conception the film consists of a unimolecular layer.

Since the film is formed spontaneously, it must offer a resistance to a decrease in area. It is possible to measure this resistance purely mechanically by compressing the film by the motion of a surface or "unidimensional" piston. If the area of the film is plotted against the force per unit length acting on the "piston," it is found that at a certain area there is an abrupt change in the force necessary to cause a decrease in area, while both larger and smaller areas correspond to a constant force per unit length. This critical area may be interpreted as the area into which the molecules in the surface can be packed together closely in a unimolecular layer. It is clear that experiments of this kind can be used to obtain information about the orientation of the molecules and their dimensions (Langmuir).

Thus if we assume that the molecules are oriented vertically in the

film, l , the length of a molecule will be given by

$$l = \frac{a}{Ad}$$

where a is the quantity of substance in the film, d its density and A the area. Further, the volume of a molecule is

$$v = \frac{M}{N_0 d}$$

where M is the molecular weight and N_0 the Avogadro number. This gives for the cross-sectional area of the molecule, l

$$l = \frac{v}{l} = \frac{M}{N_0} \frac{A}{a}.$$

The cross-sections of long-chain fatty acids calculated by this formula are found to be independent of the length of the chain, in agreement with the hypothesis of vertical orientation. The numerical values found are about $l = 22$ sq. Å, corresponding to the reasonable value of about 5 Å for the linear dimensions. For the length of the molecule we find, e.g., $l = \text{ca. } 24$ Å for palmitic acid.

The rigid orientation which characterises films of this kind gives the surface layer mechanical stability and prevents the free movement of the molecules. The law of equipartition of energy therefore cannot be applied to these molecules taken separately. In other cases the surface film has an essentially different structure in which the molecules are able to move freely in the surface although they are oriented. In this case each molecule will possess the usual mean energy $\frac{1}{2}kT$ per degree of freedom. A film of this kind is termed a gaseous film, and in some respects can be considered as a "two-dimensional gas."

If a surface piston is used as described above, it is found in these cases that F , the force per unit length on the piston, varies continuously with the area, the equation

$$FA = kT \quad (27)$$

(where A is the area occupied by a molecule in the film) being approximately satisfied. This equation is exactly analogous to the ordinary equation of state II. (2) for three-dimensional gases. If the energy is expressed in ergs k has the value 1.373×10^{-16} [III. (14)], so that if the area occupied by a molecule is expressed in sq. Å, the value of k in equation (27) will be $k = 1.373$.

It is easily seen that the quantity F in this equation is identical with the difference between γ_0 , the surface tension of the pure medium, and γ , the surface tension of the surface covered by film, i.e.,

$$F = \gamma_0 - \gamma.$$

Since F is positive, the Gibbs' adsorption equation is qualitatively confirmed.

2. COLLOID CHEMISTRY

(a) **The Colloidal State.** In the heterogeneous systems with which we have so far dealt the single portions of each phase have generally been so large that the heterogeneity of the system could be recognised without difficulty. However, if we consider for example a

suspension of solid particles in a liquid, it is clear that continued subdivision of the solid phase, so that its particles approach "molecular dimensions," is equivalent to a gradual transition from a heterogeneous to a homogeneous state. We should therefore anticipate that there will exist a state intermediate between homogeneous and heterogeneous states in which the systems dealt with will differ essentially from those which we have so far treated. We cannot of course predict *a priori* whether such systems will be sufficiently stable to realise in practice, but experience shows that they can in fact be obtained. Systems belonging to this intermediate region (referred to as the *colloidal state*) are of frequent occurrence in nature and of considerable importance.

As explained above, colloidal solutions must be regarded as intermediate between coarse suspensions or emulsions on the one hand and true solutions on the other. There is no sharp distinction between these three classes of systems, but the properties which are specially typical of colloidal systems are usually associated with particle sizes between 1 and $100\mu\mu$, where $1\mu\mu = 10^{-6}$ mm. = 10\AA . Particles which are much larger than this will be affected by gravity, and if they are heavier than the surrounding medium will sediment completely in a system at rest. If the particles are much smaller than $1\mu\mu$ the system will have the properties of a true solution. As might be expected, the properties of colloidal solutions are very dependent upon the particle size, approximating in some cases to coarse suspensions and in other cases to true solutions.

The medium in which the colloid is present is generally known as the *dispersion medium*, while the colloid itself is the *disperse phase*. The more finely divided the colloid, the greater its *degree of dispersion*, which may be defined as the ratio of the surface of the disperse phase to its volume. If the particles are spherical and of uniform size, the degree of dispersion is obviously inversely proportional to the radius of the particles.

One method of recognising colloidal solutions is by means of membranes, which can easily be prepared so as to be impermeable to colloids, but permeable to the solvent and solutes of ordinary molecular size. Colloids can therefore be separated from substances with a higher degree of dispersion (*i.e.*, a smaller particle size) by dialysis or "ultrafiltration" through membranes, *e.g.*, of collodion or animal membranes.

Colloidal solutions are also characterised by the so-called *Tyndall effect*. This depends on the fact that when light passes through a solution containing sufficiently large particles it is scattered, so that the path of the beam of light becomes visible. There is also a very weak scattering with true solutions, and in this case the wave-

length of some of the scattered light is altered in a way depending on the nature of the scattering molecule (the *Raman effect*).

We should expect colloidal solutions to obey the laws of dilute solution, but since the effect of the solute on the vapour pressure or freezing point of the solvent depends on the number of dissolved molecules, and not on their nature or size, it is clear that a given weight of substance in the colloidal state will have a much smaller effect than the same weight of substance present as molecules of normal size. In general, therefore, the vapour pressure, freezing point, etc., of a colloidal solution do not differ appreciably from the corresponding properties of the pure dispersion medium.

The most important type of colloidal solution is liquid, as assumed above. A liquid colloidal solution is often termed a "sol." Sometimes a sol will "set" to form a "gel," which is an elastic solid. Colloids are also known in which the dispersion medium is a gas. We shall not, however, deal further with the extensive terminology and classification of colloids.

(b) **The Formation of Colloidal Solutions.** In certain cases a colloidal solution is formed directly by bringing a coherent phase (a macro-phase) into contact with a solvent. This happens with many proteins, gluc, dextrin, starch, etc., in water, and with rubber in benzene, etc. In such cases we must assume that the large molecules characterising the colloidal system are already present in the macro-phase, so that the process of solution is analogous to the corresponding process for substances of low molecular weight. In other words, in these cases the macro-phase is soluble, and its molecules are so large that the solution formed is colloidal.

Colloids of this kind are termed *lyophilic colloids*, since (just as for ordinary solutions) the formation of the solution depends on the existence of forces between the medium and the solute which are at least comparable to the forces which hold together the molecules of the solute in the macro-phase. At absolute zero the relation between these forces would be the only factor determining the process of solution, but at higher temperatures the effect of the thermal agitation of the molecules will also come into play. However, the larger the molecules the less important the effect of thermal agitation in comparison with that of the intermolecular forces. This follows from the law of equipartition of energy [III. 1. b.], according to which the translational energy of a molecule (which is a measure of the effect of temperature on the tendency of the system to become homogeneous) is independent of the molecular size: on the other hand, the potential energy due to the interaction of the molecule with its surroundings will increase with increasing size. On the basis of these considerations it is seen that the solubility

relations of colloids (and of large molecules in general) will exhibit more extreme behaviour than is found for substances with lower molecular weights.

Colloidal solutions are also known of substances which in their normal state have no tendency to go into solution as described above. Colloids are known in which the disperse phase consists of a substance which when present as a macro-phase dissolves only slightly or not at all in the dispersion medium. Colloids of this kind are termed *lyophobic colloids*. In these cases the attraction between the molecules and the medium is so small that it is not able to overcome the mutual attraction between the molecules in the macro-phase, even when assisted by thermal agitation. In such cases the macro-phase can only be converted into a colloidal solution if it is first broken down to a particle size of the order of magnitude which is characteristic of colloidal solutions. Once this is effected a suspension of the finely divided material in the dispersion medium will be able to form a colloidal solution at least momentarily, but on account of the forces of cohesion between the colloid particles there will always be a tendency for them to be precipitated in the form of the original macro-phase. However, under certain conditions this type of colloidal solution may be stable for long periods, so that in practice the fundamental distinction between lyophilic and lyophobic colloids may be to some extent obscured.

The fine subdivision necessary to form a lyophobic colloid can often be effected by purely "mechanical dispersion," i.e., by grinding up the coarser particles with the dispersion medium. If the disperse phase is liquid it may be emulsified by shaking up with the dispersion medium. However, the colloidal solutions thus obtained often have a low stability.

Another method which may be used is known as "electro-dispersion," and consists in striking an electric arc between metal electrodes immersed in a solvent. This is an excellent method for preparing colloidal solutions of metals such as silver, gold and platinum in water. The solutions are strongly coloured and remarkably stable.

Colloidal solutions of this kind can also be obtained by "condensation," starting with solutions of normal molecular weight. A very wide range of processes exhibit this behaviour, but we may mention in particular that when a substance is present in solution it can often be obtained in a colloidal state by adding a solvent in which it is insoluble. Further, the product of many chemical reactions (double decompositions, oxidation or reduction reactions, etc.) is often formed as a colloid, or at least appears temporarily in the colloidal state. For example, colloidal gold is obtained by

reducing gold salts with formaldehyde, colloidal sulphur in the reaction between sodium thiosulphates and strong acids, and colloidal arsenic trisulphide by passing hydrogen sulphide into neutral aqueous solutions of arsenic trioxide.

It will be seen from the methods used for preparing lyophobic colloids that in these cases the particle size or degree of dispersion will be a very variable quantity. The same is true to some extent of lyophilic colloids, but in this case they can often be obtained from natural sources with fairly constant molecular size. If they are prepared artificially the size of their particles or molecules will usually be subject to very considerable variations. Colloidal solutions are described as *mono-disperse* or *poly-disperse*, according to whether the particles they contain are of uniform or non-uniform size.

(c) **The Thermodynamic Stability of Colloidal Solutions.** The fact that solutions of lyophilic colloids are formed spontaneously shows that they are stable with respect to the initial materials, i.e., the coherent macro-phase of the colloid plus the dispersion medium. Re-precipitation of the colloid in the form of the macro-phase can therefore only take place if the medium is changed. Such a change often brings about simultaneously a change in the chemical nature of the colloid.

For lyophobic colloids the state of affairs is quite different. Since they are not formed spontaneously from the coherent macro-phase and the medium, the colloidal state must be unstable with respect to these components under the experimental conditions. The fact that many lyophobic colloids in practice appear to be stable must depend on the presence of factors which have an inhibiting kinetic effect on the formation of the macro-phase. We shall return later to this important question.

The thermodynamic instability of a lyophobic colloid may be measured by the free energy difference between the colloid and the macro-phase, which is identical with the work which must be supplied to carry out the process of dispersion. For a given system this work will be a function of the degree of dispersion. It may be calculated on the assumption that the work involved in the dispersion process can be expressed by equation II. (39),

$$A = 4\pi r^2 \gamma, \quad (28)$$

so that the free energy changes involved in dispersing N particles is given by

$$\Delta F = 4\pi N r^2 \gamma. \quad (29)$$

Since the mass of N particles is $4/3\pi r^3 N d$, where d is the density, we

have for 1 gram of colloid

$$\Delta F = \frac{3\gamma}{rd}. \quad (30)$$

If γ is expressed in C.G.S. units (as in Table I, Chapter II), and r in cm., ΔF in this expression will be given in ergs. If ΔF is in calories and r in Å, the equation becomes

$$\Delta F = \frac{7.17\gamma}{rd}. \quad (31)$$

We see from this expression that the work of dispersion is proportional to the degree of dispersion. The work required to disperse a given quantity of substance is thus greater the smaller the particles into which it is divided. In other words, the more finely divided a colloid is, the greater will be its tendency to return to the state of the macro-phase. This result might appear to be irreconcilable with the fact that substances which have a small attraction for the solvent and thus form lyophobic colloids when dispersed nevertheless have a certain small "true" solubility, *i.e.*, they will go spontaneously into solution as single molecules, which according to the above considerations ought to be the state characterised by the greatest instability.

It should however be realised that the derivation of equations (29) to (31) is entirely static in character and does not take into account the thermal motion of the particles, which (as emphasised above) has a "dispersing" effect. In order to investigate how far thermal effect can be neglected, and for what particle sizes they begin to be of importance, it is necessary to compare the kinetic and potential energies of the particles. The work given by equation (28) must be regarded as the molecular potential energy of the particles with respect to the macro-phase, and for a reasonable mean value of γ it may be estimated as $1,000r^2$. The mean kinetic energy of each

particle is $\frac{3}{2}kT$, independent of the particle size, and is thus about 6×10^{-14} at ordinary temperatures. We thus see that for particles of colloid dimensions the thermal energy is very small compared with the potential energy, so that the above deductions about the stability of colloidal solutions are not affected. On the other hand, when the particles approach molecular dimensions, the kinetic energy becomes comparable with the potential energy. In this range, therefore, a decrease in potential energy (corresponding to a decrease in particle size) will be accompanied by an increase in the probability that the thermal energy of the particle will tear it away from the macro-phase and enable it to exist independently. The presence of a finite solubility as single molecules is thus explained.

The existence and thermodynamic stability relations of colloidal solutions also lead to more general considerations about the concept of solubility. In an ordinary solubility equilibrium we normally deal with a macro-phase (e.g., a sparingly soluble liquid) K_1 and a solution of single molecules of K_1 in the solvent K_2 . In this case stable equilibrium corresponds to the simultaneous presence of the solute in its maximum and minimum degrees of dispersion.

If the macro-phase is now dispersed as *large* suspended drops, we shall find that the solubility remains independent both of the size of the drops and of their amount. On further dispersion, however, the solubility will increase, analogously to the fact previously mentioned [II. 2. a.] that very small drops have an appreciably larger vapour pressure than large ones. As the degree of dispersion increases, we thus find that the equilibrium amount of finely dispersed solute (i.e., the dissolved molecules) increases at the expense of the coarsely dispersed solute (i.e., the suspended drops) while the degrees of dispersion of these two states approach closer together. It is clear however from considerations of symmetry that when the system changes in this way the equilibrium number of the smaller particles present cannot remain unaffected by the number of the larger particles, since the number of the latter actually does affect the number of the smaller particles present (i.e., the "solubility"). It is thus necessary to modify the concept of the solubility equilibrium to a marked degree, since according to the above considerations it must be regarded as a distribution of the matter present among different particle sizes, both the *size* of the particles and the *quantity* of them present in a given volume being of importance. The most general form of solubility equilibrium represents a distribution equilibrium for a given quantity of substance in a dispersion medium, in which particles of all sizes may in principle be present in the equilibrium state of the system. The ordinary solubility equilibrium represents a special case in which only particles of maximum and minimum sizes are present, i.e., the macro-phase and single dissolved molecules. These general considerations are of importance in understanding both heterogeneous solubility equilibria and the nature of colloidal solutions.

It must however be remembered in making use of these ideas that (as shown above) under ordinary conditions the kinetic energy is not sufficient to establish an equilibrium involving intermediate particle sizes, since with lyophobic material the equilibrium state consists almost entirely of macro-phase and molecular solute. In order to realise the more general distribution in practice it is therefore necessary to set up conditions under which the larger particles are supplied with energy from non-thermal sources, this energy increasing with the size of particle. One way of doing this is by macroscopic motion in the system, i.e., by shaking. The effect of shaking is well known and has already been mentioned as a practical method for dispersion or emulsification.

In equilibria involving single molecules mechanical motion is of no importance, since the velocities which can be obtained in practice are small compared with the velocity of molecular translation. In the case of large particles the reverse is true. The kinetic energy due to the motion of the system as a whole is proportional to the third power of the linear dimensions of the particle (supposed spherical), while the potential energy which opposes subdivision is proportional to the

second power only. Shaking will therefore readily set up mechanical dispersion of the large particles until a stationary equilibrium state is reached in which the sizes and kinetic energies of the various particles present obey a distribution law depending upon the experimental conditions. The ordinary solubility equilibrium in a system at rest represents a special case of this general equilibrium state.

The considerations of the last paragraph will not however apply in the case of "thread-shaped" molecules, of which many substances of high molecular weight are built up, *e.g.*, polystyrene and cellulose.

It is important to distinguish between the thermodynamic instability of lyophobic colloids, dealt with above, and the instability which such colloidal solutions exhibit in practice. The precipitation which takes place on keeping such solutions is known as *coagulation*, and it often consists merely in the union of the colloid particles to larger aggregates, without the formation of the coherent macro-phase we have so far dealt with. The coagulated matter thus need not possess the properties of the macro-phase: for example, a small change in conditions will often cause it to revert spontaneously to the colloidal solution, a process which is known as *peptisation*.

(d) **The Brownian Movement. The Diffusion of Colloids.** The particles in a colloidal solution are in a state of continual movement, called the Brownian movement, after its discoverer. The vigour of this movement increases with decreasing particle size, but is detectable even with particles which are so large that they are visible under an ordinary microscope. In the *ultramicroscope*, in which much smaller particles can be observed by a specially arranged illumination, the Brownian movement is a very striking phenomenon.

The Brownian movement may be regarded as the thermal molecular motion of the colloid particles. According to the law of equipartition of energy [III. 1. b.], the translational energy is the same for all particle sizes, and the velocity can therefore be calculated [III. 2. a.] according to the equation

$$mw^2 = 3kT, \quad (32)$$

where m is the mass of the particle. The relatively large value of m reduces w to an observable quantity. On account of the numerous collisions with the molecules of the dispersion medium which a particle undergoes even in a very small time interval, it is clear that the displacements which can be observed and recorded microscopically are not identical with the actual distance travelled. It is, however, easily seen that the mean values of this distance will be directly proportional to the recorded displacements.

The presence of the Brownian movement will result in the *diffusion* of the colloid from a higher to a lower concentration. This diffusion tendency may be regarded as depending on the purely

geometrical fact that when a particle moves according to the laws of chance, it is more likely to increase than to decrease its distance from a given point. If this idea is developed quantitatively, we obtain the following simple formula,

$$r^2 = 2Dt \quad (33)$$

or,

$$\frac{dr}{dt} = \frac{D}{r} \quad (34)$$

where r^2 is the mean value of the square of the displacement which a particle undergoes in time t , and D is a constant depending on the substances involved and on the temperature. It can be shown that this constant is identical with the quantity designated by the same symbol in the next paragraph. The equation can be verified by measuring the change in position of a large number of particles in given time intervals.

The laws governing the motion of single particles are of course fundamental for macroscopic diffusion in colloidal solutions. If the plane surface of a dilute colloidal solution is brought into contact with pure dispersion medium, movement of the colloid will take place in a direction at right angles to the surface between the solution and the solvent. The sharp boundary will soon disappear, being replaced by a continuous concentration fall. If the solution is dilute, the number of particles which in unit time move along the x -axis across a plane parallel to the surface is given by the diffusion equation VIII. (39), also valid for ordinary dissolved molecules, i.e.,

$$dn = DO \frac{dc}{dx} dt, \quad (35)$$

where O is the area of the plane, and the *diffusion coefficient* D is obviously the amount diffusing in unit time through unit area when the concentration gradient dc/dx is unity. The amount of substance which accumulates per unit time on account of diffusion between two parallel planes P_1 and P_2 distant dx apart can be written as Vdc , where $V = Odx$ is the volume included between the planes. It is also given by equation (35) as the difference in the amounts diffusing through P_1 and P_2 in unit time. We thus have

$$Vdc = DO \left[\left(\frac{dc}{dx} \right)_{x+dx} - \left(\frac{dc}{dx} \right)_x \right] dt,$$

or,

$$\frac{dc}{dt} = D \frac{d^2c}{dx^2}. \quad (36)$$

If the concentration gradient is constant everywhere, dc/dt will be

zero, i.e., there is a *stationary state* set up by diffusion. This is impossible in natural diffusion processes, and can only be obtained if the concentrations in two fixed planes are artificially kept constant by supplying and removing substance. In the case of ordinary molecular solutions (for which the equations are also valid) this may be realised if P_1 is the surface of a more soluble modification of the substance and P_2 the surface of a less soluble modification. The diffusion layer then consists of a "stream of substance" of constant velocity, and it is easily seen from equation (35) that the diffusion coefficient is given by the amount of substance dissolved from P_1 or deposited on P_2 per unit time and per unit area, multiplied by the distance between the surfaces and divided by the difference in the concentrations at P_1 and P_2 .

The value of the diffusion coefficient D will of course depend on the *mobility* B of the particle in the medium, i.e., the velocity with which it moves in the medium when acted on by unit force. Since in a liquid medium the resistance to motion is always large, we can write

$$u = KB = \frac{K}{G}, \quad (37)$$

where u is the velocity, K the force and G the resistance. For spherical particles which are large compared with intermolecular distances in the medium, the resistance is given by Stokes' law, II. (42), as

$$\frac{1}{B} = G = 6\pi\eta r, \quad (38)$$

where r is the radius of the sphere and η the viscosity of the medium. In order to calculate the force K_1 we use the fact that the work necessary to move the particle through a distance dx must be equal to the difference in chemical potential corresponding to this change of position. If the gas laws are valid for the solution, this law is expressed by equation III. (53),

$$K = kT \frac{d \ln c}{dx}, \quad (39)$$

whence

$$u = ktB \frac{d \ln c}{dx}. \quad (40)$$

uc is the amount of substance passing the surface in unit time, so that from the definition of the diffusion coefficient

$$uc = D \frac{dc}{dx}, \quad (41)$$

and by combining (40) and (41) we obtain

$$D = kTB = \frac{kT}{G} = \frac{k}{6\pi\eta r} \quad (42)$$

This equation may be used to determine either the Boltzmann constant k (and hence the Avogadro number), or the viscosity of the medium η , or the radius of the particle r , assuming in each case that the remaining two quantities are known.

By means of considerations already used in electrochemistry [IX. 2. n.], it is easily seen that the diffusion coefficient of a binary salt can be obtained from (42) by writing

$$B = \frac{2UV}{U + V}$$

where U and V are the mobilities of the two ions.

(e) Sedimentation Equilibria. Determination of Avogadro's Number. In the preceding sections we have treated colloid particles as molecules, and by introducing the law of the equipartition of energy we have implicitly assumed that the laws of dilute solutions are valid for colloidal solutions. An important proof of the correctness of this assumption is afforded by investigations of the sedimentation equilibria of colloids (Perrin).

If a colloid solution is in equilibrium under the force of gravity, then if the particles are sufficiently small it will not be possible under ordinary experimental conditions to detect any difference in concentration at different heights above the surface of the earth. If on the other hand the particles are large, such a difference will be detectable, and if the particles are heavier than the medium, sedimentation will take place, an equilibrium being eventually reached in which the lower portion of the solution is more concentrated than the upper portion.

This equilibrium can be calculated on the basis of Boltzmann's distribution law [III. 2. c.], according to which the following equation is valid for molecules free to move,

$$\psi = kT \ln \frac{N_B}{N_A}, \quad (43)$$

where N_B is the concentration of molecules at a level B and N_A the concentration at a level A , there being a difference of potential energy of $\psi = \psi_A - \psi_B$ per molecule between the two levels. If the level A is higher than the level B , ψ can be expressed as the work necessary to raise a particle from B to A against the force of

gravity. For a spherical particle of radius r the force acting on the particle is

$$K = \frac{4}{3}\pi r^3 g (d - d_0), \quad (44)$$

where g is the acceleration due to gravity, and d and d_0 the densities of the particle and the dispersion medium respectively. If l is the difference in height, we have

$$\psi = Kl = \frac{4}{3}\pi r^3 gl (d - d_0),$$

and by combining this with (43) we obtain

$$\ln \frac{N_B}{N_A} = \frac{4\pi g (d - d_0)}{3kT} l r^3. \quad (45)$$

If we assume $d - d_0 = 0.1$ and $r = 1\mu\mu = 10^{-7}$ cm. (which with water as dispersion medium corresponds to a molecular weight M of about 2,500), we find for $l = 10$ cm. and $T = 293$,

$$\ln \frac{N_A}{N_B} = 1.02 \times 10^{-4}.$$

The difference of concentration is thus only about 0.01%. If the radius is $10\mu\mu$ and M therefore about 2.5×10^6 (taking $d = 1$), the other values remaining unchanged, we find

$$\ln \frac{N_A}{N_B} = 0.102$$

corresponding to a concentration difference of about 10%. If $r = 100\mu\mu$, this difference of 10% will exist between levels only 0.1 mm. distant.

These calculations can be verified by counting particles (*e.g.*, suspended mastic particles) at different heights by means of the ultramicroscope. The most important source of error lies in the difficulty of preparing a homogeneous or monodisperse material, *i.e.*, a solution in which all the particles have the same size. The agreement between the experimental results and the equations derived above shows the applicability of the gas laws to colloidal solutions.

Conversely, the equations may be assumed to be valid, and the results used to determine

$$k = \frac{R}{N_0}$$

and hence Avogadro's number, N_0 . Determinations carried out by this method have given the value

$$N_0 = 6.06 \times 10^{23}$$

in very good agreement with the values found by other methods [III. 1. a.].

(f) **Sedimentation Velocity.** In dealing with the viscosity of liquids [II. 2. b.], it was mentioned that this property could be determined by measuring the rate of fall of spherical particles, using Stokes' law. The same law was also used in (d) for calculating the diffusion coefficients of colloids. The sedimentation velocity can be calculated similarly by combining (37) and (38), giving

$$u = \frac{K}{6\pi\eta r}.$$

The "force of diffusion" (39) must now, however, be replaced by the "force of sedimentation" (44). We thus obtain

$$u = \frac{2g(d - d_0)r^2}{9\eta}. \quad (46)$$

Equation (46) is of course only valid when the sedimentation process can be regarded as irreversible, i.e., when the system is so far removed from equilibrium that the quantity

$$kT \frac{d \ln c}{dx}$$

can be neglected in comparison with K . (At equilibrium this quantity is equal to K , as expressed by III. (53).) The larger and heavier the particles, the more nearly will this condition be fulfilled.

If η and $d - d_0$ are known, measurements of u can be used to determine the particle size by means of equation (46). However, for small values of r the sedimentation velocity under the influence of gravity will be much too small to be measured in practice.

If the sedimenting force is increased, the method becomes applicable in practice for much smaller particles. This principle is employed in the *ultra-centrifuge* (Svedberg), in which by using speeds of rotation up to 150,000 revolutions per minute a sedimenting force 10^8 times as great as gravity can be obtained. In this way it has proved possible to measure sedimentation velocities and equilibria even with solutes of ordinary molecular weight. The extent of sedimentation is determined by photographic measurements during rotation.

(g) **The Electric Charge of Colloids.** As mentioned above, lyophobic colloids are thermodynamically unstable. The fact that in practice many such colloidal solutions can be kept unchanged for long periods must therefore be attributed to special kinetic causes.

The apparent stability of lyophobic colloids is generally explained

by assuming that the particles carry charges of the same sign, and therefore repel each other. The electrostatic forces are "long range" forces, *i.e.*, they fall off according to an inverse square law, while the cohesive forces which bring about aggregation are (like chemical forces in general) "short-range" forces, *i.e.*, they fall off according to a much higher inverse power of the distance. It is thus easily seen how the electrostatic forces may prevent two particles from coming sufficiently close for the cohesive forces to become effective. As a result of this the colloidal solution may remain unchanged for long periods.

The presence of an electric charge on colloids may be demonstrated by observing their behaviour in an electric field. It is found in general that the particles move in the field, this phenomenon being described as *cataphoresis* or *electrophoresis*. The direction in which the particles move can be used to decide the sign of the charge. Particles of metals, sulphur, arsenic trisulphide, silicic acid, starch and mastic usually move towards the anode, and are thus negatively charged, while ferric hydroxide and other metallic hydroxides move towards the cathode, and are thus positively charged. The addition of electrolytes will often change the direction of motion. In particular, the addition of acid tends to make the charge more positive, and the addition of alkali to make it more negative. The mobility of colloid particles depends on the magnitude of the charge and the size of the particles, and hence exhibits large variations; it is, however, often of the same order of magnitude as the mobilities of ions of ordinary molecular dimensions. Thus the mobilities of colloidal particles of silver and gold (expressed as the velocity in water under a potential gradient of 1 volt/cm.) are about $2-4\mu/\text{sec.}$, while the corresponding mobility of the K^+ -ion is $6.6\mu/\text{sec.}$ [IX. 1. e.]. This behaviour shows that the colloidal particle carries a very large charge, since otherwise the large size of the particle would give it a very low velocity.

The charge on the colloid particles must of course be balanced by a charge of the same magnitude but opposite sign in the surrounding liquid, since the system as a whole is electrically neutral. This charge will normally be present as ordinary molecular ions, which may either be distributed in the solution (in which case the "colloid salt" may be said to be completely dissociated), or may be closely associated with the colloid particle.

The electric charge on the colloid can be attributed to various causes in different cases, but will always be due to the gain or loss of ions or electrons. If the colloid has acid or basic properties the charge may be produced by protolytic reaction with the solvent, a certain number of molecules in the surface taking part in the

reaction. It is thus easy to explain the positive charge of the metallic hydroxide colloids and the negative charge of silicic acid colloids, while the negative charge on colloidal arsenic trisulphide can be explained by supposing that on the surface a number of *SH*-groups take the place of *S*-atoms and are able to lose a proton. In addition to such chemical changes, adsorption phenomena are also often of importance in determining the charge on colloids. For example, in the case of a colloidal solution of silver bromide, an excess of Ag^+ -ions or Br^- -ions will produce respectively a positive and a negative charge, being adsorbed on the surface of the particles. It has been found that metallic colloids (*e.g.*, colloidal gold solutions) can only be prepared in the presence of small quantities of salts, the anion of which is adsorbed on the surface of the metal particles, giving them a negative charge. There is thus an abundance of possibilities to account for the charge on colloid particles, though opinion is often divided as to the nature of the process in particular cases. The number of charges will always be small compared with the number of atoms in the particle, so that the chemical composition of the colloid is not appreciably altered by the presence of the charge.

In the case of lyophilic colloids it is not necessary to assume the presence of an electric charge on the particles in order to account for their stability. However, many of these colloids possess acid or basic properties, and they will then undergo protolytic reactions which result in a change in their charge and which have a large effect upon their ordinary properties.

(h) **Osmotic Pressure.** Just as in the case of ordinary solutions, the osmotic pressure of a colloidal solution can be defined as the excess pressure to which the solution must be subjected in order that it may be in equilibrium (with respect to solvent) with pure solvent. In certain cases this definition can be directly applied to colloidal systems, *e.g.*, to a solution of rubber in benzene. In most cases, however, the colloid will be electrically charged and the solution will therefore contain oppositely charged ions which will have an osmotic effect depending on their concentration and osmotic coefficient. This coefficient may vary greatly from case to case. In addition to this complication due to ionisation, it is necessary to investigate further what is meant by the "solvent" in these cases, since it is possible that the semi-permeable membrane used for measurements will be permeable not only to the actual dispersion medium, but also to other substances the presence of which is associated with the colloid. In general, therefore, we are dealing with ternary systems, and the investigation of simpler systems is not of much importance in this connection. We can

predict from the laws of dilute solution that the osmotic pressure of colloidal solutions will be small compared with the osmotic pressure of ordinary solutions of the same concentration by weight, provided that the number of oppositely charged ions associated with each colloid particle is not large. As already stated, the same will of course be true of the vapour pressure lowering and freezing point depression of the solvent. These two last properties will, however, be greatly affected by the presence of small quantities of substances of low molecular weight, and are therefore not applicable for characterising the degree of dispersion or particle size of the colloid. On the other hand the osmotic method, in which the colloid is divided from a colloid-free solution by means of a semi-permeable membrane, is applicable both in principle and in practice, since we shall show below that in this case the effect of substances of low molecular weight can be definitely eliminated. We shall in fact see that under the simplest possible experimental conditions the classical laws for osmotic pressure are applicable to a colloidal solution in a binary mixed solvent.

The relation between the osmotic pressure and the concentration of non-permeating molecules has been previously derived in connection with membrane potentials [IX. 2. m.], it being assumed that the mixtures are ideal. This assumption will now be dropped. In order to derive the laws required for determining the size (*i.e.*, the molecular weight) of colloid particles we shall consider as the simplest case a ternary system consisting of a binary permeating solvent containing the two components K_1 and K_2 and a single monodisperse colloidal component K_3 . To begin with, we shall assume that all three substances are non-electrolytes. The problem is now to find the relation between the osmotic pressure and the molar concentration of K_3 in such a system, the composition of the solvent in the inner liquid being kept constant.

We shall consider a solvent of composition $n_1K_1 + n_2K_2$ to which dn_3 gram-molecules of K_3 are added. The "inner liquid" (1) therefore has the composition

$$n_1K_1 + n_2K_2 + dn_3K_3$$

where n_1 and n_2 are constant. Previous considerations [IX. 2. m.] show that at equilibrium the outer liquid (2) (containing no K_3) will contain K_1 and K_2 in a ratio differing from that in the inner liquid, and will also be under a different pressure. If the pressure on the inner liquid is kept constant and equal to zero, that on the outer liquid will be dp . Our problem is how to determine the relation between the colloid concentration and the osmotic pressure, *i.e.*, between dn_3 and dp . It is also of interest to determine dx , the difference between the compositions of the solvent in the outer and inner liquids.

We shall use the following abbreviations, where (1) and (2) refer to the inner and outer liquids respectively,

$$\left. \begin{aligned} \left(\frac{\partial \mu_{1(1)}}{\partial n_3} \right)_{p, n_1, n_2} &= \beta_{13}, \\ \left(\frac{\partial \mu_{2(1)}}{\partial n_3} \right)_{p, n_1, n_2} &= \beta_{23}, \end{aligned} \right\} \quad (47)$$

$$\left. \begin{aligned} \left(\frac{\partial \mu_{1(2)}}{\partial x} \right)_p &= \beta_{1x}, \\ \left(\frac{\partial \mu_{2(2)}}{\partial x} \right)_p &= -\beta_{2x}, \end{aligned} \right\} \quad (48)$$

and also the relations

$$\left. \begin{aligned} \left(\frac{\partial \mu_{1(2)}}{\partial p} \right)_x &= V_1, \\ \left(\frac{\partial \mu_{2(2)}}{\partial p} \right)_x &= V_2, \end{aligned} \right\} \quad (49)$$

where V_1 and V_2 are the differential molar volumes, which can be taken to be the same in (1) and (2).

The inner liquid which we are considering can be synthesised in the osmotic apparatus by adding $dn_3 K_3$ to the pure solvent $n_1 K_1 + n_2 K_2$. Since at equilibrium the potential changes for the permeating substances are the same on the two sides of the membrane, we can write

$$\left. \begin{aligned} \beta_{13} dn_3 &\approx \beta_{1x} dx + V_1 dp, \\ \beta_{23} dn_3 &\approx -\beta_{2x} dx + V_2 dp, \end{aligned} \right\} \quad (50)$$

whence by eliminating dx ,

$$\frac{dp}{dn_3} \approx \frac{\beta_{13}\beta_{2x} + \beta_{23}\beta_{1x}}{V_1\beta_{2x} + V_2\beta_{1x}}. \quad (51)$$

Application of the Gibbs-Duhem equation to the inner liquid gives

$$x_{1(2)}\beta_{1x} = x_{2(2)}\beta_{2x}, \quad (52)$$

giving on combination with (51),

$$\frac{dp}{dn_3} = \frac{\beta_{13}x_{1(2)} + \beta_{23}x_{2(2)}}{V_{(2)}}, \quad (53)$$

where $V_{(2)} = V_1x_{1(2)} + V_2x_{2(2)}$ is the molar volume of the outer liquid.

This equation is of general validity for dilute solutions. The effect of deviations from ideal behaviour is best expressed by means of activity coefficients. Introducing equation IV. (64),

$$d\mu = RTd\ln x + RTd\ln f$$

into (47), we obtain

$$\left. \begin{aligned} \beta_{13} &= \left(\frac{\partial \mu_{1(1)}}{\partial n_3} \right)_{p, n_1, n_2} = -\frac{RT}{n_{(1)}} (1 - \alpha_{13}), \\ \beta_{23} &= \left(\frac{\partial \mu_{2(1)}}{\partial n_3} \right)_{p, n_1, n_2} = -\frac{RT}{n_{(1)}} (1 - \alpha_{23}), \end{aligned} \right\} \quad (54)$$

where $n_{(1)} = n_{1(1)} + n_{2(1)}$ and

$$\left. \begin{aligned} \alpha_{13} &= n_{(1)} \left(\frac{\partial \ln f_{1(1)}}{\partial n_3} \right)_{p, n_1, n_2}, \\ \alpha_{23} &= n_{(1)} \left(\frac{\partial \ln f_{2(1)}}{\partial n_3} \right)_{p, n_1, n_2}. \end{aligned} \right\} \quad (55)$$

We now introduce these β values in (53), writing

$$\frac{dn_3}{n_{(1)} V_{(1)}} = m \quad (56)$$

and replacing $-dp$ by the osmotic pressure P this gives

$$P = RTm[1 - (\alpha_{13}x_{1(2)} + \alpha_{23}x_{2(2)})] \quad (57)$$

If α_{13} and α_{23} are zero, i.e., if the solution is ideal as regards the addition of colloid, then (57) simplifies to

$$P = RTm \quad (58)$$

Equation (56) shows that for high dilutions m is identical with the volume molarity of the solution with respect to colloid. (58) thus becomes the classical expression for osmotic pressure, which has thus been shown to apply to a mixed solvent under the conditions specified.

According to (55), α_{13} and α_{23} are identical with the α -coefficients defined by V (50). By V (60) they will therefore satisfy the condition

$$\frac{\alpha_{13}}{\alpha_{23}} = -\frac{n_{2(1)}}{n_{1(1)}},$$

and hence at sufficiently high dilutions the condition

$$\frac{\alpha_{13}}{\alpha_{23}} = -\frac{x_{2(2)}}{x_{1(2)}}$$

We thus see that under the conditions specified by V (57), the quantity $\alpha_{13}x_{1(2)} + \alpha_{23}x_{2(2)}$ is zero *in general*. This condition is, according to [V. 5], the applicability of Henry's law, which we may assume to be valid for the mixed solvent. The osmotic pressure is therefore given *in general* by the classical expression (58), even if (as is usually the case) the addition of colloid alters the activity coefficients of the components of the solvent.

In applying this method to the calculation of the osmotic pressure of colloidal solutions and hence to a determination of the particle size of the colloid, attention must be paid to the special effect which the colloidal character of the solution has on the *range* of validity of the equations. In the first place it must be noted that to neglect the number of solute molecules in comparison with the total number of molecules in the solution is a better approximation for a colloidal

solution than for an ordinary solution of the same percentage composition by weight, so that from this point of view the range of validity of equation (58) is somewhat extended. On the other hand, the volume molarity defined by (56) will differ from the volume molarity of the solution to a greater extent in the case of colloids, so that even when the mole fraction of the colloid is small it will contribute a considerable proportion of the total volume on account of the large size of the particles. This means that for moderately concentrated colloidal solutions the value of m defined by (56) must be used in (58).

As regards the determination of dx (the change in the composition of the solvent in the outer liquid caused by the addition of colloid) we shall only point out that it can of course be obtained by eliminating dp from (50), and that it will depend upon the values of the coefficients defined by equations (47) to (49). Of these, α_{11} and α_{21} must be expected to have much larger values than the corresponding coefficients for ordinary solutions, so that in the osmotic equilibrium *the composition of the solvent in the outer liquid may differ considerably from its composition in the inner liquid.*

On account of the electric charge which is generally present in colloids, the simple laws of osmotic pressure will not be expected to hold if one of the components of the solvent is a salt present at a considerable concentration. It is, however, easily shown that if the concentration of the colloid ions is small compared with the salt concentration in the solvent, then the electric membrane potential which is set up will be very small and without any appreciable effect upon the osmotic pressure. The laws derived above will therefore also be valid for electrolytic colloids provided that a strong salt solution is used as the dispersion medium.

A phenomenon related to osmotic phenomena is the swelling which takes place when many substances of high molecular weight are brought into contact with a liquid medium, *e.g.*, gelatine with water or rubber with benzene. The process consists of the penetration of the medium into the solid phase (in which it is soluble), which causes a corresponding increase in volume. The swelling can be stopped by subjecting the solid phase to a certain pressure, known as the "swelling pressure." The phenomenon is analogous to the penetration of the solvent into an osmotic cell, the tendency for the system to become homogeneous being opposed by the mechanical stability of the polymerised solid instead of by a membrane. This mechanical stability is not however usually unlimited, and the polymerised substance will eventually go into solution as a lyophilic colloid.

(i) **Coagulation and Peptisation.** It has already been stated that the chief reason for the apparent stability of lyophobic colloids lies in the presence of an electric charge upon the colloid particles. By adding an electrolyte to solutions of such colloids the state of charge of the particles can be altered, and at a suitable electrolyte concentration the charge may be completely neutralised. When this is

the case the colloid is said to be at its *iso-electric point* (Hardy). When the charge on the particles is diminished or removed the stability will also diminish, and more or less complete coagulation of the colloidal solution will take place.

The effect of added electrolyte is believed to consist in association of the colloid ion with either the cation or the anion of the added salt (depending upon the charge on the colloid). This association may be regarded as an adsorption of the ions in question upon the surface of the colloid ion. This supposition agrees with the fact that the coagulating effect increases rapidly with the valency of the coagulating ion. Thus with an arsenic trisulphide sol identical coagulating effects are found with the following molar concentrations of salts of different types,

Salt type	Molar concentration
<i>KCl</i>	50 $\times 10^{-3}$
<i>CaCl_2</i>	0.65 $\times 10^{-3}$
<i>AlCl_3</i>	0.095 $\times 10^{-3}$

On the other hand, the coagulating powers of monovalent and divalent anions of the same concentration differ very little: this shows that the As_2S_3 colloid is negatively charged, since it is the charge on the positive ion of the coagulating ion which has a large effect. If we assume (in agreement with the mechanism proposed above) that it is the amount of charge adsorbed which is the important factor, then we see from the form of the adsorption isotherm [1. a.] (assumed the same for ions of different valency), that the salt concentration necessary to produce a given coagulating effect will depend upon the valency of the ions in the way described above. In addition to this, ions of high valency will generally be adsorbed more strongly than ions of low valency, thus increasing the effect of ionic type still further.

The coagulating effect of salts can also be interpreted (without assuming any actual association process) in terms of an "ionic atmosphere" [VII. 2. c.] formed round the colloid ion and resulting in a decreased mutual electrostatic repulsion.

It should also be noted that the coagulation of lyophobic colloids is a slow process in which a continuous fall in the degree of dispersion can be observed by various methods. Thus if a small quantity of electrolyte is added to a gold sol the colour will be seen to change slowly from red to blue, corresponding to the formation of larger particles from smaller ones. With addition of increasing amounts

of salt the velocity of coagulation appears to reach a limiting value corresponding to the rate at which the discharged particles can collect together by virtue of their Brownian movement into aggregates large enough to separate out of solution as a coherent product. It has been found possible to treat this maximum coagulation velocity on the basis of ordinary kinetic principles.

As previously stated, it is often possible to *peptise* the coagulated substance by a small change in the composition of the dispersion, *i.e.*, to reconvert it into the colloidal solution from which it separated. In some cases the composition of the medium is sufficient to determine whether the colloid is present in a coagulated or a dispersed state (other conditions being kept constant), and the coagulation and peptisation are then said to be *reversible*. Such reversibility is only possible when the coagulated and the dispersed material differ only slightly from a thermodynamic point of view: this usually means that the process of coagulation is due to *adhesion* between the colloid particles caused by a diminution in their charge, the particles retaining their individuality and not forming the "macro-phase" in the ordinary sense of this term. The change of charge which brings about peptisation can be effected by a change in the electrolyte concentration: either by adding ions which are adsorbed on the particles, or by removing electrolyte so that desorption of ions takes place. Thus coagulated ferric hydroxide can be brought into solution again by washing with pure water. A sulphur sol is coagulated by the addition of sodium chloride, and is peptised again on dilution with water. In the last case the effect of salt concentration has been found to vary with the degree of dispersion (*i.e.*, the particle size), the large particles being salted out first, and by means of partial coagulation and peptisation it has proved possible to fractionate the colloid according to particle size and thus to obtain approximately mono-disperse colloidal solutions.

In some cases a small concentration of salt will cause a colloidal solution to coagulate, while further addition of salt will peptise the coagulated product. This phenomenon must be interpreted as due to a change in the sign of the charge carried by the colloid, the particles in the dilute and the concentrated salt solutions being oppositely charged.

Peptisation cannot always be regarded as a simple electrolyte effect depending only on the charge and valency of the ions, but often (like adsorption processes in general) depends upon markedly individual phenomena. For example, the silver halides are peptised by an excess of either of the salt solutions used to precipitate them: *e.g.*, silver bromide is peptised by an excess of either bromide ions or silver ions. These two ions have a very strong attraction for one

another (as shown by the low solubility of the salt), and this attraction will still be effective if the ions at the surface of the particle are not present in exactly equivalent quantities. Thus if silver ions are present in excess, they will be adsorbed by the precipitated particles which will acquire a positive charge : similarly, an excess of bromide ions will produce a negative charge.

Hydrogen and hydroxyl ions are particularly effective as peptising agents. Thus a solution of sodium hydroxide will peptise many substances which have not been formed by coagulation, but which are present as finely divided solid or liquid phases, *e.g.*, finely divided carbon or drops of turpentine. This action may be attributed to specially strong adsorption of the ions H_3O^+ and OH^- , but in many cases it can also be interpreted as an acid-base effect, in agreement with the fact that colloids often have protolytic properties. It will be easily understood that acids are usually effective towards negatively charged colloids, and alkalis towards positively charged ones.

(j) **The Solubility Relations of Lyophilic Colloids.** An exact characterisation of the coagulation and peptisation behaviour of lyophobic colloids involves many difficulties, arising partly from non-uniformity in particle size and partly from the sensitivity of the equilibrium towards the charge on the particles and other factors. In the case of lyophilic colloids, on the other hand, the solubility relations are normally simpler. This is largely due to the circumstance already dealt with that the solubility equilibrium for lyophilic colloids is *thermodynamically stable*, and not dependent upon special factors like electric charge. Further, in many cases the molecular weight (or particle size) of naturally occurring lyophilic colloids is often known ; thus hæmoglobin has a molecular weight of 16,700, egg albumin 34,000, while most proteins and naturally occurring colloidal carbohydrates have considerably higher values. In some cases, however, these colloids consist of mixtures of components of widely differing molecular weights. The problem is a particularly simple one when it is possible to isolate the solid colloid phase in a well-defined crystalline state, as is the case with egg albumin.

The solubility of lyophilic colloids is often interpreted as a special hydration effect. Such an effect would not however be a special attribute of colloids and only expresses the attraction between molecules of different species which is a necessary condition for miscibility (*cf.* IV. 3. i.). In this respect there is no distinction between the formation of hydrates with definite stoichiometric composition and formation of loose association complexes.

A special characteristic of most lyophilic colloids (*e.g.*, proteins) is

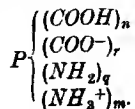
the dependence of the solubility upon the *salt concentration* and the *acidity*. Thus the solubility of crystallised egg albumin in ammonium sulphate solutions containing between 25% and 30% of salt at 18° is given by the equation

$$\log s = -0.22S + \beta$$

where S is the ammonium sulphate concentration in % and β a constant depending on the acidity of the liquid. We see that in this case the solubility is reduced to one-half by an increase of 1.37% in the salt concentration. Although the "salting-out effect" is a phenomenon exhibited by non-electrolytes in general [IV. 3. j.], the presence of a *large* effect like that illustrated above is a special colloid phenomenon.

This behaviour is readily understood in terms of the considerations advanced in (c) as to the part played by molecular size in chemical equilibria. If any substance is distributed between two media, the thermal molecular motion will tend to set up a uniform distribution between the two media. This tendency is opposed by the difference in potential energy which the substance undergoes in passing from one medium to the other. The final equilibrium state represents the resultant of these two opposing tendencies. The potential energy depends on the interaction of the particle with the surrounding medium, and for particles of similar structure may be regarded as being approximately proportional to the surface area of the particle. It is therefore clear that the potential energy difference, and hence the distribution ratio c_1/c_2 , will be very dependent on the molecular size in the case of large molecules. We must therefore anticipate that the effect of a change in medium on solubilities and distribution ratios will be much greater when the particles have colloidal dimensions than when they are of the normal molecular size.

The effect of *acidity* on the solubility of lyophilic colloids depends upon the protolytic nature of these substances. A protein is believed to contain a certain number of carboxyl groups and amino groups, so that it can be written schematically



In some cases it also contains other groups having acidic or basic properties. The molecule represented above is thus amphoteric, as it contains $n + m$ acidic groups and $p + q$ basic groups. It will therefore gain protons as the acidity increases, forming a cation, and will lose protons as the acidity decreases, forming an anion.

It is not a very simple matter to determine the acid constants of the various groups, since several of them will usually be undergoing protolytic reaction in a given range of acidity.

The "iso-electric point" is also of importance in connection with lyophilic colloids. This point is best defined in agreement with [VII. 3. e.] as the acidity at which the total charge on the colloid due to protolysis is zero. If the solid phase defining the solubility equilibrium is the electrically neutral ampholyte it will be found in agreement with [VII. 3. i.] that the solubility is a minimum at the iso-electric point.

INDEX

ABSORPTION coefficient, 144
 ——— spectrum, 347
Accumulator, 290
Acidity, 208
 ——— constant, 209, 212
 ——— exponent, 210
Activation, 251, 260
 ——— energy, 248, 258
 ———, photochemical, 260, 341
Activity coefficients, 103 ff., 186, 189
 ———, absolute, 105
Adiabatic change, 40
Adsorption, 350
 ——— density, 363
 ———, Gibbs' equation, 358
 ———, heat of, 354
 ——— indicators, 357
 ——— isotherm, 351
Affinity, 320 ff.
Allotropic change, 55, 328
Amalgam concentration cells, 287
Amorphous substances, 46
Amphiprotic, 219
Ampho-ions, 220
Ampholyte, 219
Angstrom unit, 62, 261
Anisotropic substances, 47
Anode, 266
Aprotic, 221, 357
Arrhenius, 184, 203
Atomic model, 339
 ——— p-*series*, 334
Autoprotolysis, 207, 215
Avogadro's number, 61, 354, 377
 ——— law, 64

Balmer series, 343
Basic dissociation constant, 213, 215
Bergmann, 320
Berthelot, 286, 320, 325
Berthollet, 160, 320
Bimolecular reactions, 239
Bjerrum, 186
Bohr, 339, 341
Boiling point elevation, 147
Boltzmann, 63
Boltzmann's distribution law, 72, 82,
 86, 202, 296
 ——— constant, 65
Boundary phase, 359

Boyle's law, 36, 64
Brownian movement, 374
Buffer mixtures, 211
Bunsen's absorption coefficient, 144

CALOMEL electrode, 297, 315
Capacity factor, 2
Capillary active, 361
 ——— phase, 359
Carnot, 4, 13
Carnot's cycle, 13
Catalysis, heterogeneous, 256
 ———, homogeneous, 243
 ———, protolytic, 246
Cataphoresis, 380
Cathode, 266
Cells, galvanic, 281
Chemical potential, 26 ff., 93, 98, 140
Clapeyron, 52, 57
Clark cell, 285
Clausius, 52, 57, 63
Cohesion, 42, 78
 ——— pressure, 78
Colding, 1
Collision complex, 241, 251
 ———, number, 248
Colloids, 367 ff.
Complex, critical, 253
Components, 90
Concentration, 90
 ——— cells, 286
Conductivity, electrolytic, 264 ff.
Conjugate solutions, 114
Constants, chemical, 333
Contact potentials, 278, 293
Correspondence principle, 342
Corresponding acids and bases, 204
Critical point, 51
 ——— solution temperature, 112
Cryohydric point, 117
Crystal lattice, 83
Crystalline state, 47, 83, 326
Cyclic processes, 3, 13

Debye, 88, 202, 271
Desorption, 352, 358
Deuterium, 319
Dielectric constant, 200, 275
Differential quantities, 97

- Diffusion, 74, 374
 ——— coefficient, 255, 375
 ——— potential, 310
 Dipoles, 220
 Dispersion, degree of, 368
 ——— medium, 368
 Dissociation, 165, 177, 230
 ——— constant, 202 ff.
 ———, degree of, 165, 184
 ——— energy, 346
 ——— heat of, 348
 ——— of atoms, 339
 ——— of water, 208, 215, 245
 ——— of weak electrolytes, 273
 Distribution law, *Boltzmann's*, 72, 82, 86, 202, 296
 ——— ———, *Maxwell's*, 66, 71, 85, 247
Donnan, 304
 Double layer, electrical, 278
Dulong and Petit, law of, 84, 87

Einstein, 85, 87, 260
 Electrode, gas, 302
 ——— potentials, 290
 ———, redox, 298
 ——— of the second kind, 296
 Electrodispersion, 370
 Electrolytes, strong, 186, 269
 ———, weak, 202, 273
 Electrolytic conduction, 264
 Electromotive force, 275, 281
 ——— potential, 279
 Electron, 256, 264, 278, 291, 298, 336
 Electrophoresis, 380
 Energy, 1
 ——— of activation, 248, 258
 ——— equipartition of, 62
 ——— equipotential, 5
 ——— internal, 37
 ——— kinetic, 63 ff.
 ——— potential, 5, 72, 81, 84
 ——— quantum of, 76
 ——— rotational, 63, 345
 ——— thermal, 11
 ——— translational, 63 ff.
 ——— units of, 2
 ——— vibrational, 63, 246
 Enolisation, 169
 Enthalpy, 24
 Entropy, 18 ff.
 Equilibrium conditions, 23, 25
 ——— chemical, 180
 ——— homo-heterogeneous, 170
 ——— photochemical, 262
 Equipartition of energy, 62
 Erg, 2, 343
 Esters, hydrolysis of, 244
 Eutectic point, 117
 Excited states, 339
 Extinction coefficient, 260

Faraday's law, 265, 282
Fick's law, 255
 Fluid state, 47
Franck, 336
 Freedom, degrees of, 62, 75
 Free energy, 24
 Freezing mixture, 118
 Freezing point depression, 151
 Frequency, 88, 260, 341
Freundlich's adsorption isotherm, 351
 Fusion, 115 ff.
 ———, heat of, 54

 Gas constant, 36
 ——— degeneration, 329
 ——— electrodes, 302
 ——— equilibria, 165, 332
 ——— laws, 36 ff., 185
 ——— reactions, 165
 ——— ———, unimolecular, 251
 ——— thermometer scale, 39
Gay Lussac, 37
Gibbs, 23
Gibbs's adsorption law, 358
 ——— phase rule, 137
Gibbs-Duhem equation, 99, 104, 106, 143, 158, 193, 289, 324, 383
 Ground state, 339
Guldberg, 160, 320

Hardy, 386
 Heat, 4 ff.
 ——— conduction of, 15
 ——— content, 24
 ——— engines, 12
 ——— latent, 10, 45
 ——— mechanical equivalent of, 40
 ——— theorem, 327
Helmholtz, 24, 283, 289
Henry's law, 144 ff., 192, 354
Hertz, 336
Hittorf, 267
van't Hoff, 154, 321
van't Hoff factor, 184, 199
Huckel, 202, 271
 Hydration, 147
 Hydrogen electrode, 287, 315
 Hydrogen ion catalysis, 243
 ——— concentration, 315
 ——— exponent, 211
 Hydroxonium ion, 206

 IDEAL mixtures, 100
 Indicators, 222, 301, 357

- Inertia, moment of, 340
 Inhibitors, 254
 Integral quantities, 97
 Intensity factors, 2
 Internal pressure, 78
 Ionic association, 202
 — atmosphere, 189
 — exchange, 357
 — reactions, 241
 — strength, 191
 Ionisation potential, 337
 —, thermal, 338
 Irreversibility, 8, 20
 Isoelectric point, 222, 386, 390
 Isomerisation equilibria, 165
 Isotropic state, 47
- Joule*, 1, 37
- Kelvin*, 14, 286
Kirchhoff's law, 33
Konowalow's rule, 107, 127, 361
- Lambert-Beer* absorption law, 260
Langmuir, 366
Langmuir's adsorption isotherm, 353
 Lattice energy, 83
Le Chatelier's principle, 116, 164, 175
Lewis, G. N., 24
Loschmidt's number, 61
- MACRO-PHASE, 359, 373
 Mass action, law of, 160, 170
 Mass, reduced, 248
Maxwell, 66, 247
Maxwell's distribution law, 66, 71, 85, 247
Mayer, J. R., 1, 40
 Melting point, 53, 55
 Membrane equilibria, 304
 — potentials, 304
 Membranes, semipermeable, 310
Millikan, 61
 Miscibility, 111
 Mixed crystals, 126
 Mixing, differential, 95
 — heat of, 92, 110
 — integral, 95
 — work of, 91, 94
 Mixtures, 89 ff.
 — ideal, 101
 — ternary, 157, 382
- Mobility, 266, 376
- Nernst*, 310
Nernst's heat theorem, 327
- ORIENTED films, 366
 Oscillator, harmonic, 84, 339, 346
 Osmotic coefficient, 192
 — pressure, 140, 153, 381
 — work, 153
Ostwald's dilution law, 204
 — solubility coefficient, 144
 Over-voltage, 318
 Oxidation potential, 300
 Oxidation-reduction electrodes, 298
- PARTIAL pressures, 90, 105
 Peptisation, 385
 Permutite, 357
Perrin, 377
 Phase equilibria, 57
 — rule, 137
 Photochemical reactions, 259
 Photoelectric effect, 341
 Photosensitisation, 345
Planck, 76, 327
Planck's constant, 88, 260, 339, 341
Poisson, 202
 Polarisation, 318
 Polarity, 111
 Polymorphism, 55
 Potential, chemical, 26 ff., 93, 98, 140
 — electric, 275
 — electrochemical, 276
 Protolysis, 204 ff.
 — constants, 213, 232
 Protolytes, 204
 Proton, 204
- QUANTUM mechanics, 335
 — theory, 76, 334
 Quinhydrone electrode, 317
- RADIOACTIVE disintegration, 250
Raman effect, 369
Raoult's law, 145 ff., 192
 Reaction, order of, 233
 Redox electrodes, 298
 — indicators, 301
 Reduction potentials, 300
 Relaxation time, 270
 Resonance potentials, 337
 Reversibility, 8, 281
 Rotational energy, 63, 345
 — spectra, 345
Rydberg's constant, 342
- SALT effects, 226
 — hydrates, 121
 Salting out effect, 115, 389
 Sedimentation, 377

- Solution, heat of, 201
 Solutions, colloidal, 368
 —, conjugate, 114
 —, dilute, 140 ff.
 —, ideal, 100
 —, solid, 126
 Solvation, 147, 272
 Sorption, 350
Sørensen, 210
 Specific heat, 31, 39, 74, 84, 176, 201
 Standard cells, 285
 — electrodes, 305
 — potentials, 294
 — states, 26
Stokes' law, 46, 273, 376, 379
 Strength exponent, 219
 Surface chemistry, 349 ff.
 — concentration, 363
 — films, 366
 — tension, 42, 360
 — work, 44
Svedberg, 379

 TAMMANN, 55
 Thermodynamic functions, 17 ff., 40,
 57, 91, 330
 — equation of state, 34
 — potential, 24 ff.
 — temperature scale, 14, 39
 Thermodynamics, first law of, 1
 —, second law of, 4
 —, third law of, 327
Thomsen, Julius, 320
 Thomson, W., 55, 286
 Titration, 224
 Transition point, 59
 Transport numbers, 267

 Triple point, 59
Trouton's rule, 83
Tyndall effect, 368

 ULTRACENTRIFUGE, 379
 Ultramicroscope, 374
 Unimolecular films, 352, 366
 — reactions, 237

 VAPORISATION, heat of, 48, 81
 Vapour pressure curves, 51, 105, 124,
 134
 — lowering, 145
 — of salt hydrates, 131
 — of saturated solutions,
 123
 Velocity constants, 234
 Viscosity, 46
 Volt-electron, 337

Waage, 160, 320
van der Waals' equation, 79
 Wavelength, 261
Weston cell, 285
 Work of ionisation, 336
 —, maximum, 11
 —, osmotic, 155

 X-RAYS, 341

 ZERO order reaction, 257

